EPITHELIAL CELL GROWTH INHIBITORS

This application is a continuation of international application number PCT/US00/16900, filed 19 June 2000, pending, which claims the benefit of U.S. Provisional Patent Application No.60/139,995, filed June 18, 1999, the disclosure of which is incorporated in its entirety.

Field of the Invention

This invention relates to a family of epithelial cell growth inhibitors useful in the diagnosis and treatment of epithelial cell cancers.

Background of the Invention

Epithelial cell cancers, for example, prostate cancer, breast cancer, colon cancer, lung cancer, pancreatic cancer, ovarian cancer, cancer of the spleen, testicular cancer, cancer of the thymus, etc., are diseases characterized by abnormal, accelerated growth of epithelial cells. This accelerated growth initially causes a tumor to form. Eventually, metastasis to different organ sites can also occur. Although progress has been made in the diagnosis and treatment of various cancers, these diseases still result in significant mortality.

The treatment of cancer is greatly enhanced by early detection. However, there are difficulties in detecting the disease in its early stages. For example, epithelial tissue-containing organs such as the prostate, ovary, and others, are not easily palpated. The detection of abnormal tumor growth in such organs is difficult without frequent screening and

appropriate markers. A substantial drawback of available cancer diagnostic assays is a high rate of false positive and negative results, making the available tests less reliable than desired. For this reason, there is a great need to identify new diagnostic as well as new therapeutic agents to improve diagnosis and treatment of cancer, for example, prostate cancer, breast cancer, colon cancer, lung cancer, pancreatic cancer, ovarian cancer, cancer of the spleen, testicular cancer, cancer of the thymus, etc.,

A novel, specific, mammary cell growth inhibitor, Mammastatin, has recently been identified and characterized. Mammastatin has been expressed from variant clones, MammA (PCT/US97/18026, ATCC# 97451, deposited 22 February 1996); MammB (PCT/US97/27147, ATCC#_____, deposited 15 June 2000); and MammC, described in copending PCT application No. PCT/US00/_____, filed on even date herewith (ATCC# _____, deposited 15 June 2000).

Mammastatin is produced and secreted by normal mammary cells, and is detected in blood samples of normal individuals. Blood concentrations of the mammary cell growth inhibitor, and particularly of the active, phosphorylated form of Mammastatin, are reduced or absent in breast cancer patients. Administration of protein comprising active Mammastatin (secreted from normal human breast cancer cells) is effective to reduce tumor size and number, and to prevent tumor growth in late stage cancer patients.

Epithelial cell growth inhibitors having similarity to Mammastatin have now been discovered, isolated, and characterized. These inhibitors bear partial sequence identity to Mammastatin at the 5' end of the sequence, and have little or no identity at the 3' end of the molecule. Like Mammastatin, the newly discovered family of epithelial cell growth inhibitors (ECGI) are differentially expressed in normal epithelial cell tissues, but not in cancerous

epithelial cell tissues. Also, like Mammastatin, the newly discovered family of epithelial cell growth inhibitors are detected in blood samples taken from normal individuals, but not in the blood of patients with epithelial cell cancers, as shown in the Examples below.

Summary of the Invention

A family of epithelial cell growth inhibitors (ECGI) have now been identified in a number of different epithelial cells. These ECGI are differentially expressed in normal epithelial cells, but not in epithelial cancer cells. As shown in the Examples below, Mammastatin-like ECGI proteins have been discovered in a variety of epithelial cell tissues, including prostate, colon, ovary, lung, spleen, testis, thymus, and others.

The ECGI of the invention are expressed in normal epithelial cells but not in cancerous epithelial cells. The Mammastatin-like ECGI proteins are encoded by nucleic acid sequences that hybridize to nucleic acid sequences encoding Mammastatin. The ECGI proteins also bind anti-Mammastatin antibody. A nucleic acid sequence encoding ECGI in prostate cells (PRT-6, SEQ ID NO: 4) has been isolated and characterized (PRT-6, ATCC#_____, deposited 15 June 2000), as described in the Examples below.

Because the ECGI of the invention are differently expressed by normal epithelial cells and not by cancerous epithelial cells, the presence or amount of the ECGI can be analyzed to diagnose cancer and/or to monitor treatment. The inventive ECGI proteins and nucleic acids encoding them also provide useful therapeutic agents to inhibit epithelial cell growth, prevent tumor formation, and treat cancer.

Brief Description of the Figures

Figure 1A is a schematic diagram of an mRNA test panel showing locations of specific tissue mRNAs for analysis.

Figure 1B is a computer scanned image of a Northern blot showing hybridization of Mammastatin nucleic acid sequence to mRNA from a variety of tissues according to the plan shown in Figure 1A.

Figure 2 is a computer scanned image of a dot blot assay showing control, Mammastatin standard protein, serum samples from breast cancer patients, and conditioned medium from normal and cancerous human prostate cells probed with anti-Mammastatin antibody, 7G6.

Figure 3 is a computer scanned image of a Western blot assay, showing normal human mammary cell lysate (A), human prostate cancer LnCap cell lysate (B), MCF7 breast cancer cell lysate (C), and normal human prostate cell lysate (D) probed with anti-Mammastatin antibody, 7G6.

Figure 4 is a computer scanned image of a Western blot assay, showing cell lysates from normal prostate cells (A), LnCap prostate cancer cells (B), normal colon cells (C), and colon cancer cells (D) probed with anti-Mammastatin antibody, 7G6.

Figure 5 is a computer scanned image of a Western blot assay, showing cell lysates from human ovarian cancer cells (B), normal human ovarian cells (C), and normal human mammary cells (D) probed with anti-Mammastatin antibody, 7G6. Lane A contained molecular weight standards.

Figure 6 is a computer scanned image of a dot blot assay showing serum samples from healthy male adults (A,C,D) and from a prostate cancer patient (B) probed with anti-Mammastatin antibody, 7G6.

Figure 7 is a computer scanned image of a DNA gel containing putative prostate ECGF DNA clones.

Figure 8 is a diagramatic representation of Prostate ECGI and its structural relationship to other sequences.

Detailed Description of the Invention

Proteins of the invention:

"Epithelial cell growth inhibitor (ECGI) proteins" of the invention are defined herein to mean Mammastatin-like proteins produced by and active to inhibit the growth of normal epithelial cells. Active, inhibitory ECGI proteins of the invention are reduced or absent in cancerous epithelial cells. The ECGI protein family disclosed herein appears to include inhibitors that are specific to each epithelial tissue, with little or no inhibitory activity across tissue types. As discussed more fully below, it is postulated that each ECGI protein contains a growth inhibitory domain and a tissue-specificity domain.

The ECGI proteins of the invention exhibit significant homology to Mammastatin, a mammary cell growth inhibitor produced by normal human mammary cells, and previously demonstrated be useful in the diagnosis and treatment of breast cancer (PCT/US97/18026). ECGI proteins bind one or more anti-Mammastatin antibodies such as 7G6 (Neomarkers, Freemont, CA), and are encoded by nucleic acid sequences sharing significant homology with nucleic acid sequences encoding Mammastatin.

Studies reported in the Examples below demonstrate the differential expression of ECGI proteins in normal epithelial cell tissues, but not in cancerous epithelial cell tissues, including breast, prostate, ovary, and colon. Like Mammastatin, the ECGI proteins of the invention appear, for example, in Western blots, as doublets or triplet bands, with one major band and one or two smaller, less prominent bands. This pattern of expression was demonstrated for Mammastatin to be due to phosphorylation of the protein. Mammastatin has an approximate molecular weight of 53 kilodaltons when phosphorylated at two sites. Smaller sized Mammastatin, 49 and 44 kilodaltons, correspond to one or none of the sites being phosphorylated. Phosphorylation of the Mammastatin protein is correlated with its inhibitory activity.

Western blots of ECGI probed with the anti-Mammastatin antibody 7G6, demonstrate the approximate size of ECGI produced by various epithelial cell tissues. As shown more fully in the Examples below (see, for example, Figures 4-5), ECGI from prostate cells migrates in a Western blot to approximately 55 kilodaltons, with less prominent, smaller bands at 51 and 46 kilodaltons suggestive of phosphorylated forms similar to the pattern seen for Mammastatin. ECGI from colon cells migrates to approximately 50 KD, with less prominent bands at approximately 47 and 43 kilodaltons. ECGI from ovarian cells migrates to approximately 60 kilodaltons.

Nucleic Acid Sequences Encoding ECGI

Nucleic acid sequences of the invention are defined herein as those nucleic acid sequences that encode ECGI proteins, as defined above. Nucleic acid sequences encoding ECGI proteins share significant sequence homology to nucleic acid sequences encoding

Mammastatin, and hybridize to nucleic acid sequences encoding Mammastatin under conditions of high stringency.

Mammastatin-like epithelial cell growth inhibitors preferably have substantial identity (at least 90%, and preferably at least 95% identity) over approximately 1000 contiguous nucleotides of a nucleic acid sequence encoding Mammastatin. Nucleic acids encoding Mammastatin include those DNA inserts of MammA (PCT/US97/18026, ATCC# 97451, deposited 22 February 1996); MammB (PCT/US97/27147, ATCC#_____, deposited 15 June 2000); and MammC, described herein (ATCC# _____, deposited 15 June 2000). Consensus sequences determined for known Mammastatin clones are shown in the Comparative Sequence Table 5 below, and as SEQ ID NO: 1 (MammA); SEQ ID NO: 2 (MammB); SEQ ID NO: 3 (MammC). Prostate ECGI nucleic acid sequence (SEQ ID NO: 4) is shown in Tables 1, 2, and 5.

ECGI can be amplified from a specific epithelial cell nucleic acid library, for example, using internal Mammastatin primers and/or by hybridization to Mammastatin under conditions of strict stringency. As shown more fully in the Examples below, nucleic acid sequences hybridizing to Mammastatin have been demonstrated in numerous epithelial tissues, including central nervous system, heart, small intestine, large intestine, appendix, rectum, lymphatic cells, bone marrow cells, lung and air passages, bladder, uterus, prostate, testis, ovary, liver, pancreas, adrenal gland, salivary gland, and mammary gland (See Figure 1).

The nucleic acid sequence of a ECGI isolated from prostate cells, for example, shares greater than 95% identity to Mammastatin at the 5' half of the molecule, with little or no identity of sequence, however, at the 3' half. It is postulated that the 5' end, sharing identity

with Mammastatin, includes a growth inhibitory domain of the molecule, whereas the 3' end, having little identity to Mammastatin, includes a tissue-specificity domain.

Diagnostic Methods

The invention further provides an *in vitro* assay for detecting active, inhibitory ECGI in patient samples, including tissues, cells, and fluids. Epithelial cell cancer and advancing metastatic disease is diagnosed by correlating the presence and type of ECGI protein in a patient's sample with that of normal or cancerous human epithelial cells. A patient's blood or tissue sample is analyzed for the ECGI protein, e.g., for the abundance of the ECGI protein and/or for its molecular weight forms. As discussed below, the absence or loss of ECGI protein, particularly of the higher molecular weight, phosphorylated forms, is correlated with a specific epithelial cell indicative of advancing metastatic disease.

Analysis of ECGI can be performed using a variety of known analytical tools and methods, including immunoassays, hybridization, PCR techniques, and the like. Preferred are immunoassay, including ELISA, Western Blot, and dot-blot analysis of a patient's sample methods, using anti-ECGI antibodies. Preferably, recombinant ECGI standards are used to provide a standard curve for reliable quantitation of inhibitor levels. Such immunoassays are exemplified by the dot-blot assays and Western blot assays shown in the examples below. In an alternative preferred embodiment of the invention, tissue samples, such as tumor biopsies, are analyzed by immunohistochemistry, or by culturing a patient's tumor cells and examining the cultures for expression of ECGI.

In a particularly preferred embodiment, an assay for the diagnosis of an epithelial cell cancer includes at least two specific antibodies: an antibody to identify the sampled tissue as

epithelial tissue, such as an anti-cytokeratin antibody, and a specific anti-ECGI antibody. For example, using an immunoblot format, prostate tissue suspected of containing the prostate cancer cells is homogenized, separated on an SDS/PAGE gel, transferred to membrane, and probed with both anti-keratin and anti-prostate ECGI antibodies. Isotype specific second antibodies that are conjugated to a suitable marker system such as peroxidase or alkaline phosphates are used to detect bound antibodies. Membranes containing bound first and second antibodies are then developed using known colormetric or fluorometric techniques and quantitated by known methods.

In the most preferred embodiment, the sample is analyzed for the size and/or phosphorylated forms of the ECGI, such as by Western Blot, using anti-ECGI antibodies. A decline or absence of the high molecular weight ECGI protein form correlates with advancing cancer.

Diagnostic kits of the invention include ECGI protein or nucleic acid sequences encoding ECGI, for example, as controls. Optionally, the diagnostic kit contains one or more antibodies that bind the epithelial cell ECGI to be detected or quantified. The antibodies may bind a Mammastatin-like domain (for example, 7G6), or may be tissue-specific ECGI antibodies. Alternatively, the diagnostic kit includes one or more amplification primer or hybridization probe for the amplification and/or detection of nucleic acid sequences encoding an epithelial cell ECGI, for example, the primers used in the Examples below.

Therapeutic Use

ECGI protein for therapeutic use is produced from epithelial cell cultures under serum free conditions or by recombinant means. Preferably, ECGI protein is produced in yeast or

higher eucaryotic cells to achieve phosphorylation of the protein. Recombinant protein is produced in host cells or by synthetic means.

Functional ECGI is administered to patients by known method for the administration of phosphoprotein, preferably by injection, to increase inhibitor levels in the bloodstream and increase the inhibitor's interactions with the desired epithelial.

The protein may be delivered to the patient by methods known in the field for delivery of phosphorylated protein agents. In general, the inhibitor is mixed with the delivery vehicle and administered by injection.

The dosage of inhibitor to be administered may be determined by one skilled in the art, and will vary with the type of treatment modality and extent of disease. Since Mammastatin inhibits approximately 50% of mammary cancer cell growth at a concentration of 10 ng/ml and stops growth at about 20-25 ng/ml *in vitro*, a useful therapeutic dosage range of ECGI is about 2.5 μg to about 250 μg administered daily dose. Preferred is approximately 125 μg daily administered dose. The aim of the administration is to result in a final body dose that is in the physiological (e.g. 15-50 ng/ml) or slightly higher range (for example, 25-75 ng/ml). For clinical use, the preferred dosage range is about 500 ng/ml for initial treatment of metastatic disease, followed by a maintenance dosage of about 50 ng/ml. In clinical studies using Mammastatin, an administered daily dose of about 50 ng/ml to about 750 ng/ml was sufficient to induce remission to Stage IV breast cancer patients.

Since active ECGI is a phosphorylated protein, it is anticipated that multiple doses of the inhibitor will be required to maintain growth inhibiting levels of ECGI in the patient's blood. Also, since ECGI generally acts as a cytostatic agent rather than a cytocidal agent, it is

expected that a maximum effect of the inhibitor will require regular maintenance of inhibitor levels in epithelial cell cancer patients.

In its preferred use, the ECGI is administered in high dosages (>50 ng/ml, preferably about 50-500 ng/ml) to induce tumor regression. Lower, maintenance doses (<50 ng/ml, preferably 20-50 ng/ml) are used to prevent cancer cell growth.

Clinical experience with administered Mammastatin in Stage IV breast cancer patients indicates a useful dose is that which maintains physiological levels of Mammastatin in the blood. Administration is preferably daily, but, may be, for example, by continuous infusion, by slow release depot, or by injection once every 2-3 days. Anecdotal evidence suggests continuous administration may induce feedback inhibition, thus, a preferred administration scheme is to administer daily dose of Mammastatin for approximately 25-28 days, followed by 2-5 days without administration.

Diagnostic Assay

Assays of the present invention for detecting the presence of the functional inhibitor in human tissue and serum are useful in screening patients for epithelial cell cancer, for screening the population for those at high risk of developing epithelial cell cancer, for detecting early onset of epithelial cancer, and for monitoring patient levels of inhibitor during treatment. For example, analysis of a patient's blood ECGI, for example, may indicate a reduced amount of high molecular weight, phosphorylated prostate ECGI, as compared with a normal control or with the patient's prior prostate ECGI profile. Such a change is correlated with increased risk of prostate cancer, with early onset of prostate cancer, and with advancing metastatic prostate cancer. Diagnostic assay for phosphorylated, active, 55 kD prostate ECGI preferably is by

Western blot immunoassay, or ELISA using specific anti-ECGI antibodies. Screening, for example, in serum, is preferably by immunoassay, e.g., ELISA, Western blot, or dot blot assay.

For best results, the patient samples should be assayed within a short time of sampling (within one week), stored at 4°C (less than one year), or frozen for long term storage. Most preferably, samples are frozen until time of assay.

EXAMPLES

The invention may be better understood by reference to the following Examples, which are not intended to limit the invention in any way.

EXAMPLE 1

Multiple Tissue Expression of ECGI

Northern blot analysis was performed on a multiple tissue expression array (Clonetech, Inc. #7775-1) to demonstrate the expression of ECGI in a variety of epithelial cell tissues. A digoxin-labeled EcoR1 fragment of Mammastatin, containing approximately 1800 base pairs of the 3' region of pMammC, SEQ ID NO: 3 (approximately nucleotide 359 - end) was used as a probe. The DIG-labeled Mammastatin cDNA was hybridized to the array in 10 ml easy HYB solution (Roche) for 16 hours at 65° C, with 65° C washes, anti-DIG antibody hybridization and CSPD development performed according to the manufacture's instructions. The blot was then exposed to Kodak X-OMAT film for 30 minutes at room temperature.

The tissue plan of the multiple tissue expression array is shown in Figure 1A. Hybridization of the Mammastatin cDNA to the mRNA of the array is shown in Figure 1B,

and demonstrates the variety of epithelial cell tissues expressing a Mammastatin-like ECGI sequence. Specific tissues that hybridized to the Mammastatin cDNA included: central nervous system, heart, small intestine, large intestine, appendix, rectum, lymphatic cells, bone marrow cells, lung and air passages, bladder, uterus, prostate, testis, ovary, liver, pancreas, adrenal gland, salivary gland, and mammary gland.

EXAMPLE 2

Normal Versus Cancerous Prostate Cells

Normal prostate cells obtained from surgical samples and cancerous prostate cells, LnCap, obtained from the American Type Culture Collection (ATCC) were incubated and analyzed for the production of a prostate ECGI. The cells were cultured in DMEM/F12 media with 40 μM calcium, supplemented with 5% Chelex-treated horse serum, 10 ng/mL EGF, 10 μg/mL insulin, 100 ng/mL Cholera toxin and 1 μg/mL hydrocortisone for four days. Conditioned media samples were then collected and analyzed.

Normal human mammary cells obtained from patient samples were incubated in the same medium and Mammastatin secreted into the culture medium was used as a control. Serum obtained from breast cancer patients was also analyzed and used as a control.

Sample fluids were collected and loaded by suction onto a nitrocellulose membrane on a dot blot apparatus. The membranes were then probed with the anti-Mammastatin antibody 7G6, and antibody binding was detected with goat-anti mouse antibody labeled with alkaline phosphates. Color was developed with NBT/BCIP substrate system (Life Technologies). The results are shown in Figure 2.

The anti-Mammastatin antibody recognized a protein produced by normal prostate cells but not cancerous prostate cells. This is analgous to the antibody's recognition of the mammary cell growth inhibitor, Mammastatin, produced by normal mammary cells, but not breast cancer cells. This data, in combination with the data from Example 1, demonstrates the production of Mammastatin-like ECGI in other epithelial cell tissues, and particularly, in prostate cells.

EXAMPLE 3

Differential Expression of ECGI in Prostate, Colon, and Ovary

Prostate

Normal prostate cells (Clonetech, Inc.), LnCap prostate cancer cells (A.T.C.C.), MCF7 breast cancer cells (A.T.C.C.) and normal human mammary cells (obtained from hospital tissue) were incubated as described above for Example 2. After at least 48 hours incubation, cells were lysed in sample loading buffer and analyzed for the presence of ECGI by Western blot, using the anti-Mammastatin antibody, 7G6 as a probe. Normal human mammary cell protein (NHMC) lysate (1 mg/ml) was used as a Mammastatin control (A). The data are shown in Figure 3.

Normal prostate cell lysate (D) contained a protein that was recognized by anti-Mammastatin antibody, while prostate cancer cells (LnCap) (B) and breast cancer cells (MCF7) (C) did not. The protein recognized in the prostate cell lysate (D) was of a similar size to that of Mammastatin (A).

Colon and Prostate

Normal prostate cells (Clonetech, Inc.), LnCap prostate cancer cells (A.T.C.C.), Sw 948 colon cancer cells (A.T.C.C.), and normal colon epithelial cells (obtained from patient surgery tissue) were incubated as described above for Example 2. Cell lysates were prepared in sample loading buffer and analyzed for expression of ECGI by Western blot, using the anti-Mammastatin antibody, 7G6 as a probe.

As shown in Figure 4, normal prostate (A) and normal colon (C) epithelial cells expressed a protein that was recognized by the anti-Mammastatin antibody, while cancer cells from these tissues did not (B,D). The differential expression of protein is similar to that demonstrated for Mammastatin in breast tissue. In addition, the pattern of bands shown in the Western blot for normal prostate and colon tissues is similar to the Phosphorylation pattern demonstrated for Mammastatin produced in normal human mammary cells. A larger prominent band is shown together with two smaller, fainter bands. This pattern has been correlated with Phosphorylation of Mammastatin.

Prostate ECGI is shown in the Western blot analysis (Figure 4) to have an approximate molecular weight of 51 kilodaltons; Colon ECGI is shown to have an approximate molecular weight of 50 kilodaltons.

Ovary

OvCar-ovarian cancer cells (A.T.C.C.), normal human ovarian cells (patient surgery tissue) and normal human mammary cells (patient surgery tissue) were incubated as described above for Example 2. After an incubation period of at least 48 hours, direct lysates were prepared by removing growth media and rinsing cells with saline and SDS-PAGE sample loading buffer until viscous. Lysates were collected and separated on 10% SDS-PAGE,

transferred electrophoretically onto nitrocellulose, and probed with the 7G6 anti-Mammastatin antibody. The data are shown in Figure 5, where lane A contains molecular weight standards; B, OvCar-ovarian cancer cell lysate; C, normal human ovarian cell lysate; and D, normal human mammary cell lysate.

Figure 5 demonstrates that a Mammastatin-like ECGI protein is produced in normal human ovarian tissues and is recognized by anti-Mammastatin antibody. The protein is not expressed in the ovarian cancer cells analyzed. The ovarian ECGI has an approximate molecular weight of 60 kilodaltons.

Example 4

Differential Detection of Prostate ECGI in Blood

Serum samples from three healthy male volunteers were analyzed for the presence of the prostate ECGI, and compared with that of serum from a prostate cancer patient. Serum samples were loaded at 400 microliter and 200 microliter samples in duplicate. The samples were drawn onto nitrocellulose by vacuum in a 96 well dot blot apparatus. The filters were then probed with the anti-Mammastatin antibody, 7G6, and developed with NBT/BCIP substrate. The data are shown in Figure 6.

Normal human mammary cell (NHMC) cultures produced standard conditioned medium for comparison. Standards, in duplicate, contained 400, 200, 100, 50, 25, 12, and 6 microliters of NHCM medium. Serum samples from healthy adult males (A,C,D) and from an adult prostate cancer patient (B) were assayed using 400 and 200 microlites of serum sample. A prominent signal from normal serum (A,C,D) demonstrated the presence of prostate ECGI, while the prostate cancer patient's serum showed only a weak signal.

Example 5

Inhibitory Activity of Prostate ECGI

Normal prostate cells (Clonetech, Inc.), PC3 and LnCap prostate cancer cells (A.T.C.C.) were plated at a density of 5.0 x 10^4 cells per milliliter in 12 well plates in RPMI medium containing 10% fetal bovine serum. After 24 hours, the cultures were supplemented with 10% conditioned medium. Each sample was run in triplicate. Plates were allowed to incubate for six days at 37°C and 5% CO_2, and at the end of the incubation period, cells were lysed with Cetrimide and counted using a Colter Counter. Percent inhibition was calculated by comparing treated versus non-treated wells, and the data shown in the table below. Androgen-insensitive PC3 cells were not inhibited by the normal prostate cell media or by the conditioned medium obtained from normal prostate cells. In contrast, LnCap cells were inhibited by the addition of growth medium, with the inhibition somewhat greater by media derived from normal prostate versus media derived from cancer cells.

Cell Type	% Inhibition by Normal Prostate medium	% Inhibition by Prostate Tumor medium
LnCap #1	22.5 +/- 3.3	8.3 +/- 0.4
LnCap #2	22.7 +/- 0.6	16.7 +/- 15.8
PC3	0	0

Example 6

Isolation and Characterization of Prostate ECGI DNA

Nucleic acid libraries were produced from the mRNA of normal prostate cells (patient surgery tissue) and from LnCap, prostate tumor cells (A.T.C.C.).

The nucleic acid sequences in the normal and cancerous prostate cell libraries were incorporated into vectors and used to transform bacteria. Colonies of bacteria expressing the normal and cancer prostate cell nucleic acid sequences were screened by hybridization with a digoxin-labeled Mammastatin nucleic acid probe under stringent conditions, as described above.

The positive colonies were selected and grown in LB broth. Plasmids obtained from the positive colonies were purified and digested with ECO R1 and Xhol to release the CDNA inserts. The digested DNA was then separated on a 1% agarose gel (see Figure 7A) and the separated DNA was subjected to Southern blot analysis using the digoxin-labeled Mammastatin fragment as a probe. As shown in Figure 7 below, two prostate ECGI clones were isolated, each having an approximate size of 2 Kb: One clone was isolated from the normal prostate tissue library (PRN2.1) and one from the LnCap prostate tumor cell library (PRT-6).

PRT-6 was further characterized, and its nucleic acid sequence was determined. As shown below in Table 1, the nucleic acid sequence encoding Prostate ECGI has substantial identity to Mammastatin (greater than 90%) at the 5' end of the molecule (approximately nucleotides 15-1032 of MammC), with little or no identity at the 3' end of the molecule. These regions of similarity and distinction are shown diagrammatically in Figure 8.

Example 7

Isolation and Characterization of Prostate ECGI DNA

Nucleic acid libraries were constructed from the mRNA or normal prostate cells (obtained from patient surgery tissue) and from LnCap prostate tumor cells (A.T.C.C.). The library cDNA was used to transfer E.coli and plated out for colony hybridization. The

colonies were screened with a digoxin-labeled Mammastatin C fragment generated by PCR using external PCR primers M200 and M2200.

[Sequence ID NO: 5] M200: GCGCCGGCCGGGCGCGACCCG

[Sequence ID NO: 6] M2200: GCAATCTCAGCGCACTGCTGC

Bacterial colonies expressing prostate ECGI clones were hybridized to the labeled Mammastatin probe under strict hybridization conditions, as described above.

Example 8

Homology of Prostate ECGI

The prostate ECGI sequence was analyzed against nucleic acid sequences present in GenBank. Portions of two molecules showed some similarity to domains within the prostate ECGI sequence: 28SmRNA and Hip55.

28SmRNA homology has been identified in many gene sequences with importance in growth regulation (Hu et al., 1999, PNAS 96:1339-1344; Mauro et al., 1997, PNAS 94:422-427). Hip55 is a protein that binds to hematopoetic progenitor type 1 kinase, a protein involved in the src signal transduction pathway (Ensena et al, 1999, JBC 274:33945-50).

Using the open reading frame known for Hip55, a putative amino acid sequence was deducted for the prostate clone. As shown below in Table 3, the translation includes several internal stop codons.

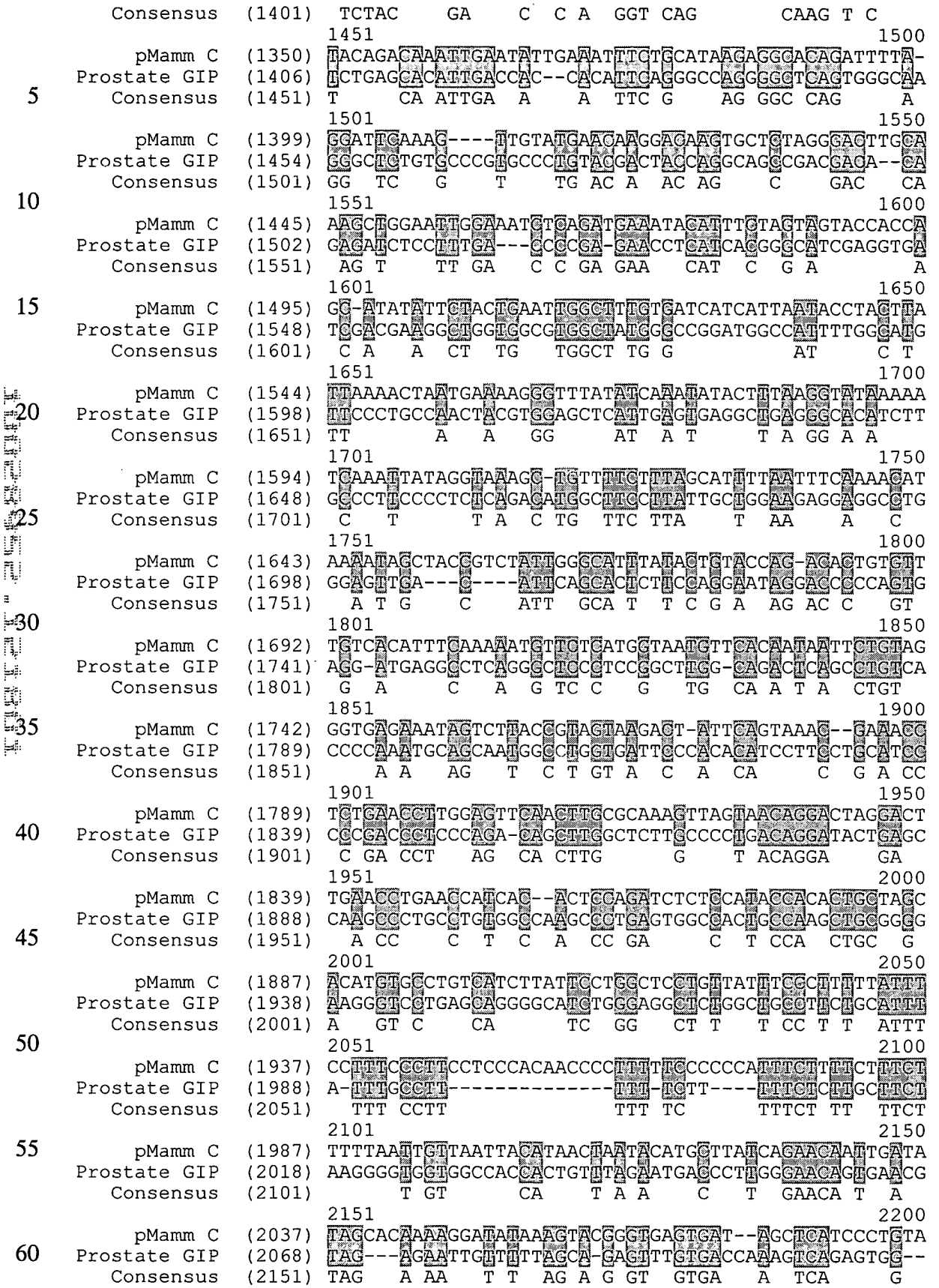
Also using the Hip55 ORF, a putative amino acid sequence was deduced for MammB and MammC sequences, shown in Tables 4 and 5.

Table 1

pMammC and Prostate ECGI

```
                          1                                                  50
     pMamm C       (1)    --------------------------------------------------
Prostate GIP       (1)    GCACGAGATTCCCACTGTCCCTACCTACTATCCAGCGAAACCACAGCCAA
   Consensus       (1)

                         51                                                 100
     pMamm C       (1)    --------------------------------------------------GA
Prostate GIP      (51)    GGGAACGGGCTTGGCGGAATCAGCGGGGAAAGAAGACCCTGTTGAGCTTG
   Consensus      (51)

                        101                                                 150
     pMamm C       (3)    ATTCGGCACGAGCACGGTGAAGAGACATGAGAGGTGTAGAATAAGTGGGA
Prostate GIP     (101)    ACTGTAGTGTGGCACGGTGAAGAGACATGAGAGGTGTAGAATAAGTGGGA
   Consensus     (101)    A TC    C  GCACGGTGAAGAGACATGAGAGGTGTAGAATAAGTGGGA

                        151                                                 200
     pMamm C       (53)    GGCCCCCGGCGCCCCCCCGGTGTCCCCGCGAGGGGCCCGGGGCGGGGTCC
Prostate GIP     (151)    GGCCCCCGGCGCCCCCCCGGTGTCCCCGCGAGGGGCCCGGGGCGGGGTCC
   Consensus     (151)    GGCCCCCGGCGCCCCCCCGGTGTCCCCGCGAGGGGCCCGGGGCGGGGTCC

                        201                                                 250
     pMamm C       (103)    GCCGGCCCTGCGGGCCGCCGGTGAAATACCACTACTCTGATCGTTTTTTC
Prostate GIP     (201)    GCCGGCCCTGCGGGCCGCCGGTGAAATACCACTACTCTGATCGTTTTTTC
   Consensus     (201)    GCCGGCCCTGCGGGCCGCCGGTGAAATACCACTACTCTGATCGTTTTTTC

                        251                                                 300
     pMamm C       (153)    ACTGACCCGGTGAGGCGGGGGGGCGAGCCCCGAGGGGCTCTCGCTTCTGG
Prostate GIP     (251)    ACTGACCCGGTGAGGCGGGGGGGCGAGCCCCGAGGGGCTCTCGCTTCTGG
   Consensus     (251)    ACTGACCCGGTGAGGCGGGGGGGCGAGCCCCGAGGGGCTCTCGCTTCTGG

                        301                                                 350
     pMamm C       (203)    CGCCAAGCGCCCGGCCGCGCGCCGGCCGGGCGCGACCCGCTCCGGGGACA
Prostate GIP     (301)    CGCCAAGCGCCCGGCCGCGCGCCGGCCGGGCGCGACCCGCTCCGGGGACA
   Consensus     (301)    CGCCAAGCGCCCGGCCGCGCGCCGGCCGGGCGCGACCCGCTCCGGGGACA

                        351                                                 400
     pMamm C       (253)    GTGCCAGGTGGGGAGTTTGACTGGGGCGGTACACCTGTCAAACGGTAACG
Prostate GIP     (351)    GTGCCAGGTGGGGAGTTTGACTGGGGCGGTACACCTGTCAAACGGTAACG
   Consensus     (351)    GTGCCAGGTGGGGAGTTTGACTGGGGCGGTACACCTGTCAAACGGTAACG

                        401                                                 450
     pMamm C       (303)    CAGGTGTCCTAAGGCGAGCTCAGGGAGGACAGAAACCTCCCGTGGAGCAG
Prostate GIP     (401)    CAGGTGTCCTAAGGCGAGCTCAGGGAGGACAGAAACCTCCCGTGGAGCAG
   Consensus     (401)    CAGGTGTCCTAAGGCGAGCTCAGGGAGGACAGAAACCTCCCGTGGAGCAG

                        451                                                 500
     pMamm C       (353)    AAGGGCAAAAGCTCGCTTGATCTTGATTTTCAGTACGAATACAGACCGTG
Prostate GIP     (451)    AAGGGCAAAAGCTCGCTTGATCTTGATTTTCAGTACGAATACAGACCGTG
   Consensus     (451)    AAGGGCAAAAGCTCGCTTGATCTTGATTTTCAGTACGAATACAGACCGTG

                        501                                                 550
     pMamm C       (403)    AAAGCGGGGCCTCACGATCCTTCTGACCTTTTGGGTTTTAAGCAGGAGGT
Prostate GIP     (501)    AAAGCGGGGCCTCACGATCCTTCTGACCTTTTGGGTTTTAAGCAGGAGGT
   Consensus     (501)    AAAGCGGGGCCTCACGATCCTTCTGACCTTTTGGGTTTTAAGCAGGAGGT

                        551                                                 600
     pMamm C       (453)    GTCAGAAAAGTTACCACAGGGATAACTGGCTTGTGGCGGCCAAGCGTTCA
Prostate GIP     (551)    GTCAGAAAAGTTACCACAGGGATAACTGGCTTGTGGCGGCCAAGCGTTCA
   Consensus     (551)    GTCAGAAAAGTTACCACAGGGATAACTGGCTTGTGGCGGCCAAGCGTTCA

                        601                                                 650
     pMamm C       (503)    TAGCGACGTCGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCATTGTG
Prostate GIP     (601)    TAGCGACGTCGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCATTGTG
   Consensus     (601)    TAGCGACGTCGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCATTGTG

                        651                                                 700
     pMamm C       (553)    AAGCAGAATTCACCAAGCGTTGGATTGTTCACCCACTAATAGGGAACGTG
```

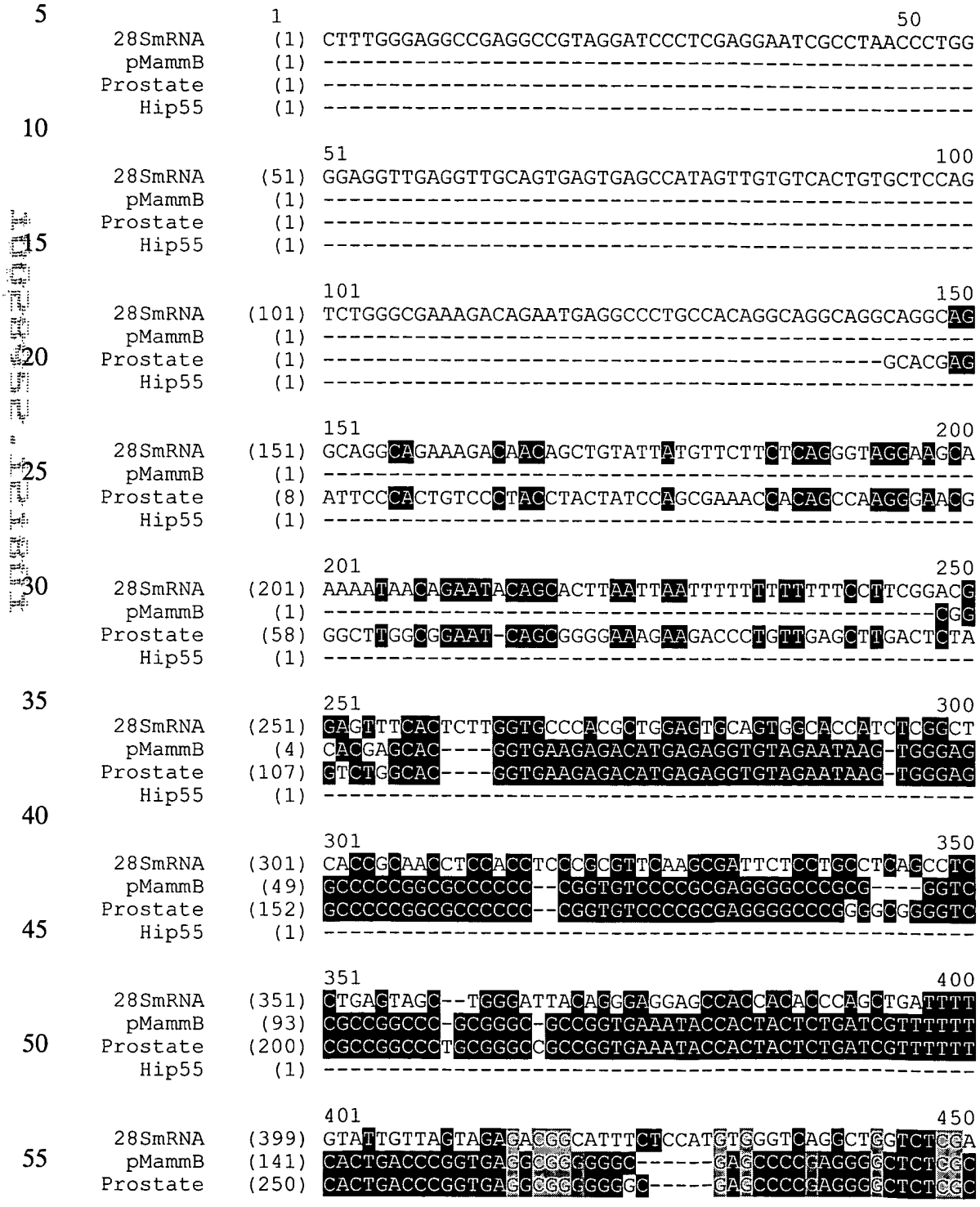
```
Prostate GIP    (651) AAGCAGAATTCACCAAGCGTTGGATTGTTCACCCACTAATAGGGAACGTG
   Consensus    (651) AAGCAGAATTCACCAAGCGTTGGATTGTTCACCCACTAATAGGGAACGTG
                      701                                             750
     pMamm C    (603) AGCTGGGTTTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACTGATGAT
Prostate GIP    (701) AGCTGGGATTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACTGATGAT
   Consensus    (701) AGCTGGG TTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACTGATGAT
                      751                                             800
     pMamm C    (653) GTGTTGTTGCCATGGTAATCCTGCTCAGTACGAGAGGAACCGCAGGTTCA
Prostate GIP    (751) GTGTTGTTGCCATGGTAATCCTGCTCAGTACGAGAGGAACCGCAGGTTCA
   Consensus    (751) GTGTTGTTGCCATGGTAATCCTGCTCAGTACGAGAGGAACCGCAGGTTCA
                      801                                             850
     pMamm C    (703) GACATTTGGTGTATGTGCTTGGCTGAGGAGCCAATGGGGCGAAGCTACCA
Prostate GIP    (801) GACATTTGGTGTATGTGCTTGGCTGAGGAGCCAATGGGGCGAAGCTACCA
   Consensus    (801) GACATTTGGTGTATGTGCTTGGCTGAGGAGCCAATGGGGCGAAGCTACCA
                      851                                             900
     pMamm C    (753) TCTGTGGGATTATGACTGAACGCCTCTAAGTCAGAATCCCGCCCAGGCGG
Prostate GIP    (851) TCTGTGGGATTATGACTGAACGCCTCTAAGTCAGAATCCCGCCCAGGCGG
   Consensus    (851) TCTGTGGGATTATGACTGAACGCCTCTAAGTCAGAATCCCGCCCAGGCGG
                      901                                             950
     pMamm C    (803) AACGATACGGCAGCGCCGCGGAGCCTCGGTTGGCCTCGGATAGCCGGTCC
Prostate GIP    (901) AACGATACGGCAGCGCCGCGGAGCCTCGGTTGGCCTCGGATAGCCGGTCC
   Consensus    (901) AACGATACGGCAGCGCCGCGGAGCCTCGGTTGGCCTCGGATAGCCGGTCC
                      951                                            1000
     pMamm C    (853) CCCGCCTGTCCCCGCCGGCGGGCGGCCCCCCCCCTCCACGCGCCCCGCG
Prostate GIP    (951) CCCGCCTGTCCCCGCCGGCGGGCCGCCCCCCC--CTCCACGCGCCCCGCG
   Consensus    (951) CCCGCCTGTCCCCGCCGGCGGGCCGCCCCCCC  CTCCACGCGCCCCGCG
                     1001                                            1050
     pMamm C    (903) CGCGCGGGAGGGCGCGTGCCCCGCCGCGCGCCGGGACCGGGGTCCGGTGC
Prostate GIP    (999) CGCGCGGGAGGGCGCGTGCCCCGCCGCGCGCCGGGACCGGGGTCCGGTGC
   Consensus   (1001) CGCGCGGGAGGGCGCGTGCCCCGCCGCGCGCCGGGACCGGGGTCCGGTGC
                     1051                                            1100
     pMamm C    (953) GGAGTGCCCTTCGTCCTGGGAAACGGGGCGCGGCCGGAAAGGCGGCCGCC
Prostate GIP   (1049) GGAGTGCCCTTCGTCCTGGGAAACGGGGCGCGGCCGGAAAGGCGGCCGCC
   Consensus   (1051) GGAGTGCCCTTCGTCCTGGGAAACGGGGCGCGGCCGGAAAGGCGGCCGCC
                     1101                                            1150
     pMamm C   (1003) CCCTCGCCCGTCACGCACCGCACGTTCGTGCT---CGTGCCGAATTCGGC
Prostate GIP   (1099) CCCTCGCCCGTCACGCACCGCACGTTCGTGGGGAACCTGGCGC-TAAACC
   Consensus   (1101) CCCTCGCCCGTCACGCACCGCACGTTCGTG     C TG CG  T    C
                     1151                                            1200
     pMamm C   (1050) ACGAGTAGCACCATTGACAATAGAGATACAAGTGCATGTATCTTTATGATT
Prostate GIP   (1148) ACCTCCATCTCCAGTCCTCA--GCGTGGCAAGCTGAGG-AGCCCTTCCT
   Consensus   (1151) AC    A C CCA TC   A  G C   CAAG   A G A C    T  T
                     1201                                            1250
     pMamm C   (1100) ATAATGAATTGTTTTCCTTTGGGTAGATATCCAGTAGTGGGATTGCTAGA
Prostate GIP   (1195) GCA--GAAG-CAGCTCACCCAACCAGAGACCCACT-------TTGGCAGA
   Consensus   (1201)   A  GAA  C   TC        AGA A CCA T       TTG  AGA
                     1251                                            1300
     pMamm C   (1150) TCACCTGGTAGTTCTATTTCTGGTTTATTGAGAAATCTTCATACTGATTT
Prostate GIP   (1235) GAGCCAGCTGCTGGCATCTCAAGGCCCAGGGCAGATCTCCCTGCTGAG--
   Consensus   (1251)    CC G T  T C AT TC  G      G  A ATCT C T CTGA
                     1301                                            1350
     pMamm C   (1200) CCATAGAGGTTGTACAAATTTACATGCCTACCAAGTGATTTTTTTAAATA
Prostate GIP   (1283) -----GAGCCGGCGCC------CAGCACTCGTCCATGTCTGGTGCAGGCA
   Consensus   (1301)      GAG   G  C       CA C CT C    TG  T  T  A   A
                     1351                                            1400
     pMamm C   (1250) TGAAAGAATGGTCTGGAGAAATGCCCCTCATTAGTATCCCCCTTTTACCT
Prostate GIP   (1322) GAAGAGGAGGCTGTGTATGAG-GAACCTCCAGAGCAGGAG------ACCT
   Consensus   (1351)   A AG A G T TG A  A  G  CCTC   AG A          ACCT
                     1401                                            1450
     pMamm C   (1300) CTCTACTGCAGAATGACTTCAAGGGGTACAGGTATTTACAAGTTTCATTA
Prostate GIP   (1365) -TCTAC----GAGCAGCCCCACTGGTGCAGCAG----CAAGGTGGTGGC
```

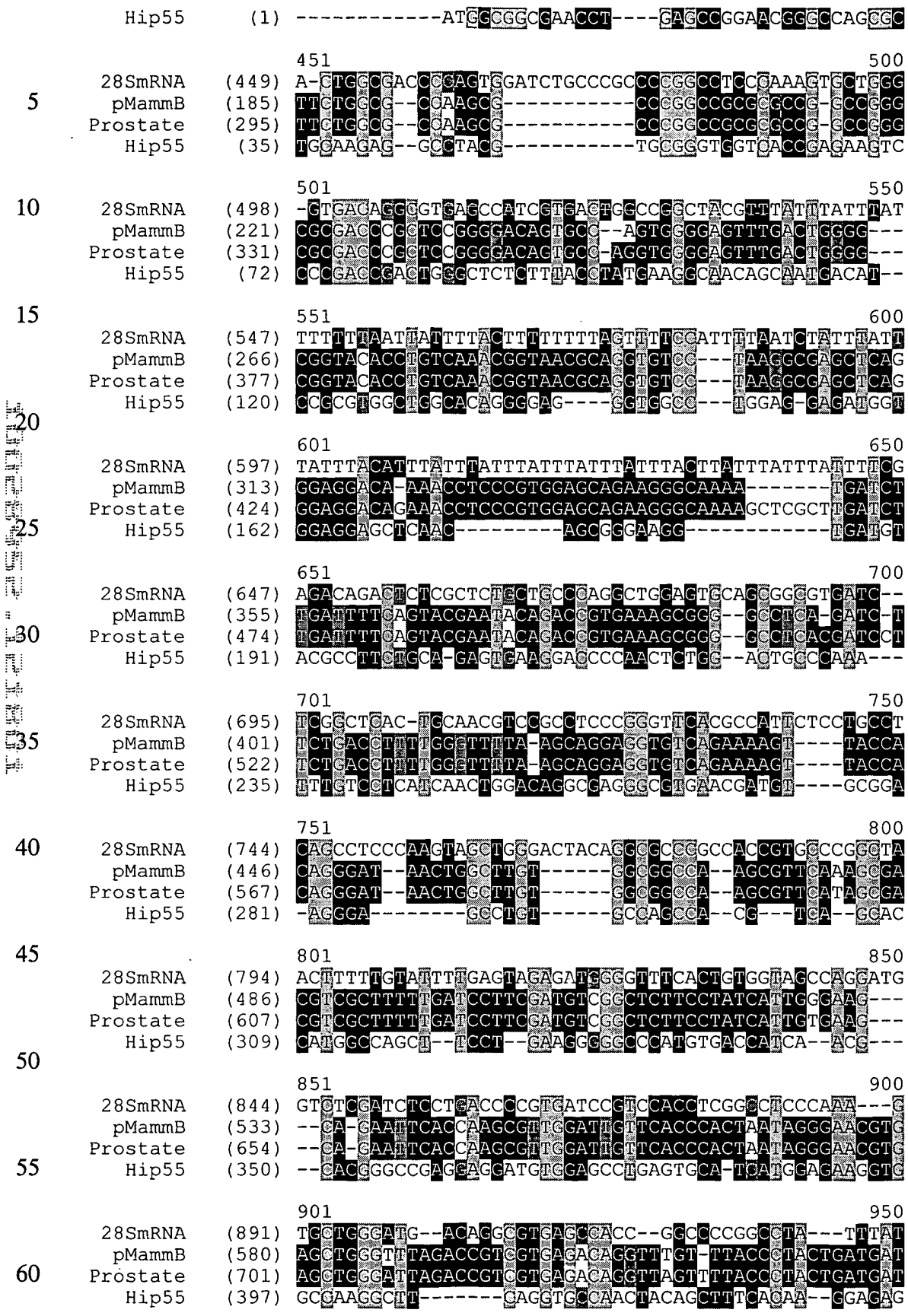

Consensus (1401) TCTAC GA C C A GGT CAG CAAG T C
1451 1500
pMamm C (1350) TACAGACAAATTGAATATTGAAATTTCTGCATAAGAGGCACAGATTTTA-
Prostate GIP (1406) TCTGAGCACATTGACCAC--CACATTCAGGGCCAGGGGGTCAGTGGGCAA
Consensus (1451) T CA ATTGA A A TTC G AG GGC CAG A
1501 1550
pMamm C (1399) GGATTCAAAG----TTGTATGAACAAGGACAAGTGCTGTAGGGAGTTGCA
Prostate GIP (1454) GGGCTCTGTGCCCGTGCCCTGTACGACTACCAGGCAGCGACGACA--CA
Consensus (1501) GG TC G T TG AC A AC AG C GAC CA
1551 1600
pMamm C (1445) AAGCTGGAATTGGAAATGTCAGATGAAATACATTTGTAGTAGTACCACCA
Prostate GIP (1502) GAGATCTCCTTTGA---GCCCGA-GAACCTCATCACGGGCATCGAGGTGA
Consensus (1551) AG T TT GA C C GA GAA CAT C G A A
1601 1650
pMamm C (1495) GG-ATATATTCTACTGAATTGGCTTTGTGATCATCATTAATACCTACTTA
Prostate GIP (1548) TCGACGAAGGCTGGTGGCGTGGCTATGGCCGGATGGCCATTTTGGCATG
Consensus (1601) C A A CT TG TGGCT TG G AT C T
1651 1700
pMamm C (1544) TTAAAACTAATGAAAAGGGTTTATATCAAATATACTTTAAGGTATAAAAA
Prostate GIP (1598) TTCCCTGCCAACTACGTGGAGCTCATTGAGTGAGGCTGAGGGCACATCTT
Consensus (1651) TT A A GG AT A T T A GG A A
1701 1750
pMamm C (1594) TCAAATTATAGGTAAAGC-TGTTTTCTTTAGCATTTTAATTTCAAAACAT
Prostate GIP (1648) GCCCTTCCCCTCTCAGACATGGCTTCCTTATTGCTGGAAGAGGAGGCCTG
Consensus (1701) C T T A C TG TTC TTA T AA A C
1751 1800
pMamm C (1643) AAAATAGCTACCGTCTATTGGGCATTTATACTGTACCAG-ACAGTGTGTT
Prostate GIP (1698) GGAGTTGA---C----ATTCAGCACTCTTCCAGGAATAGGACCCCCAGTG
Consensus (1751) A T G C ATT GCA T T C G A AG AC C GT
1801 1850
pMamm C (1692) TGTCACATTTCAAAAATGTTCTCATGGTAATGTTCACAATAATTCTGTAG
Prostate GIP (1741) AGG-ATGAGGCCTCAGGGCTCCCTCCGGCTTGG-CAGACTCAGCCTGTCA
Consensus (1801) G A C A G TC C G TG CA A T A CTGT
1851 1900
pMamm C (1742) GGTGAGAAATAGTCTTACCGTAGTAAGACT-ATTCAGTAAAC--GAAACC
Prostate GIP (1789) CCCCAAATGCAGCAATGGCCTGGTGATTCCCACACATCCTTCCTGCATCC
Consensus (1851) A A AG T C T GT A C A CA C G A CC
1901 1950
pMamm C (1789) TCTGAACCTTGGAGTTCAACTTGCGCAAAGTTAGTAACAGGACTAGGACT
Prostate GIP (1839) CCCGACCCTCCCAGA-CAGCTTGGCTCTTGCCCCTGACAGGATACTGAGC
Consensus (1901) C GA CCT AG CA CTTG G T ACAGGA GA
1951 2000
pMamm C (1839) TGAACCTGAACCATCAC--ACTCCAGATCTCTCCATACCACACTGCTAGC
Prostate GIP (1888) CAAGCCCTGCCTGTGGCCAAGCCCTGAGTGGCCACTGCCAAGCTGCGGGG
Consensus (1951) A CC C T C A CC GA C T CCA CTGC G
2001 2050
pMamm C (1887) ACATGTGCCTGTCATCTTATTCCTGGCTCCTGTTATTTCGCTTTTTATTT
Prostate GIP (1938) AAGGGTCCTGAGCAGGGGCATGTGGGAGGCTCTGGCTGCCTTCTGCATTT
Consensus (2001) A GT C CA TC GG CT T T CC T T ATTT
2051 2100
pMamm C (1937) CCTTTCCCTTCCTCCCACAACCCCTTTTTCCCCCCATTTCTTTTCTTTCT
Prostate GIP (1988) A-TTTGCCTT--------------TTT-TCTT----TTTCTCTTGCTTCT
Consensus (2051) TTT CCTT TTT TC TTTCT TT TTCT
2101 2150
pMamm C (1987) TTTTAATTGTTAATTACATAACTAATACATGCTTATCAGAACAATTGATA
Prostate GIP (2018) AAGGGGTGGTGGCCACCACTGTTTAGAATGACCTTGGGAACAGTGAACG
Consensus (2101) T GT CA T A A C T GAACA T A
2151 2200
pMamm C (2037) TAGCACAAAAGGATATAAAGTACGGGTGAGTGAT--AGCTCATCCCTGTA
Prostate GIP (2068) TAG---AGAATTGTTTTAGCA-GAGTTTGTGACCAAAGTCAGAGTGG--
Consensus (2151) TAG A AA T T AG A G GT GTGA A TCA G

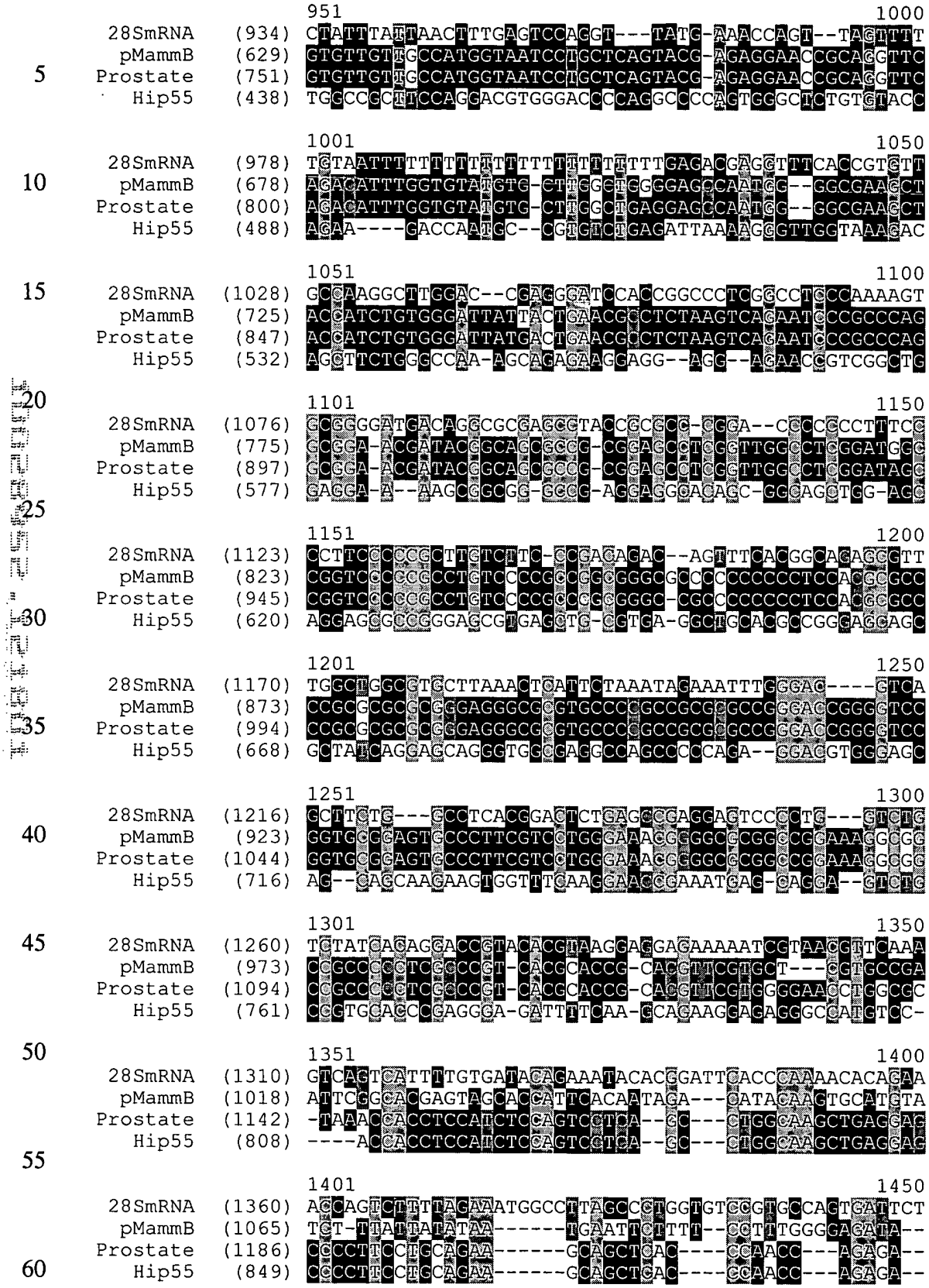

2201 2250
pMamm C (2085) ATGCTAGCACTTTGGAAGGCCAAGGCAGGCAGATCACTTGAGTGCCAGAGT
Prostate GIP (2112) ATCATGGTGGTTTGGCAG--CAGGGAATTTGTCTGTTGGAGCCT---GC
Consensus (2201) ATC T G TTTGG AG CA GG A T T GAG C G
2251 2300
pMamm C (2135) TGGAGACCAGCCTGGGCAACATGGTGAAACCCTGTCTCTAGAAAAAATA
Prostate GIP (2157) TCTGTGGTCCCCACTCCATTTCTCTGTCCGTCTGCCTGGGCTATGGGAAG
Consensus (2251) TC C CC CA TG C CTG CT C A A
2301 2350
pMamm C (2185) CAAAAATTTAGCCGGGCGTGCTGGCACACACCTGTAGTCTCAGCTACTCT
Prostate GIP (2207) TGGGGATGCAGATGGCCAAGCTCCCAC---CCTGGGTATTCAAAAAC---
Consensus (2301) AT AG GG C GCT CAC CCTG TCA AC
2351 2400
pMamm C (2235) GAGGGCTGAGGTGGGAAGATTGATTGAGCCCAGGAGGTGGAAGCTGCAGC
Prostate GIP (2251) ---GGCAGACACAACATG-TTCCTCCACGCGGCTCACTCGATGC--CTGC
Consensus (2351) GGC GA A G TT T A C T GA GC C GC
2401 2450
pMamm C (2285) AGTGGGCTGAGATTGCGCCATTGCACTCCAGCCTGGGTGAGAGAGAGAGA
Prostate GIP (2295) AGGCCCCAGTGTGTGCCTCAACTGATTCTGACTTCAGGAAAAGTAAAA-A
Consensus (2401) AG C C G G TGC CA A TC C T G A AG A A A
2451 2498
pMamm C (2335) CCCTGTCTCAAAAAAAAAAA----------
Prostate GIP (2344) A--------AAAAAAAAAAAACTCGAGAAGCTTTGGACTTCTTCGCCA
Consensus (2451) AAAAAAAAAAA

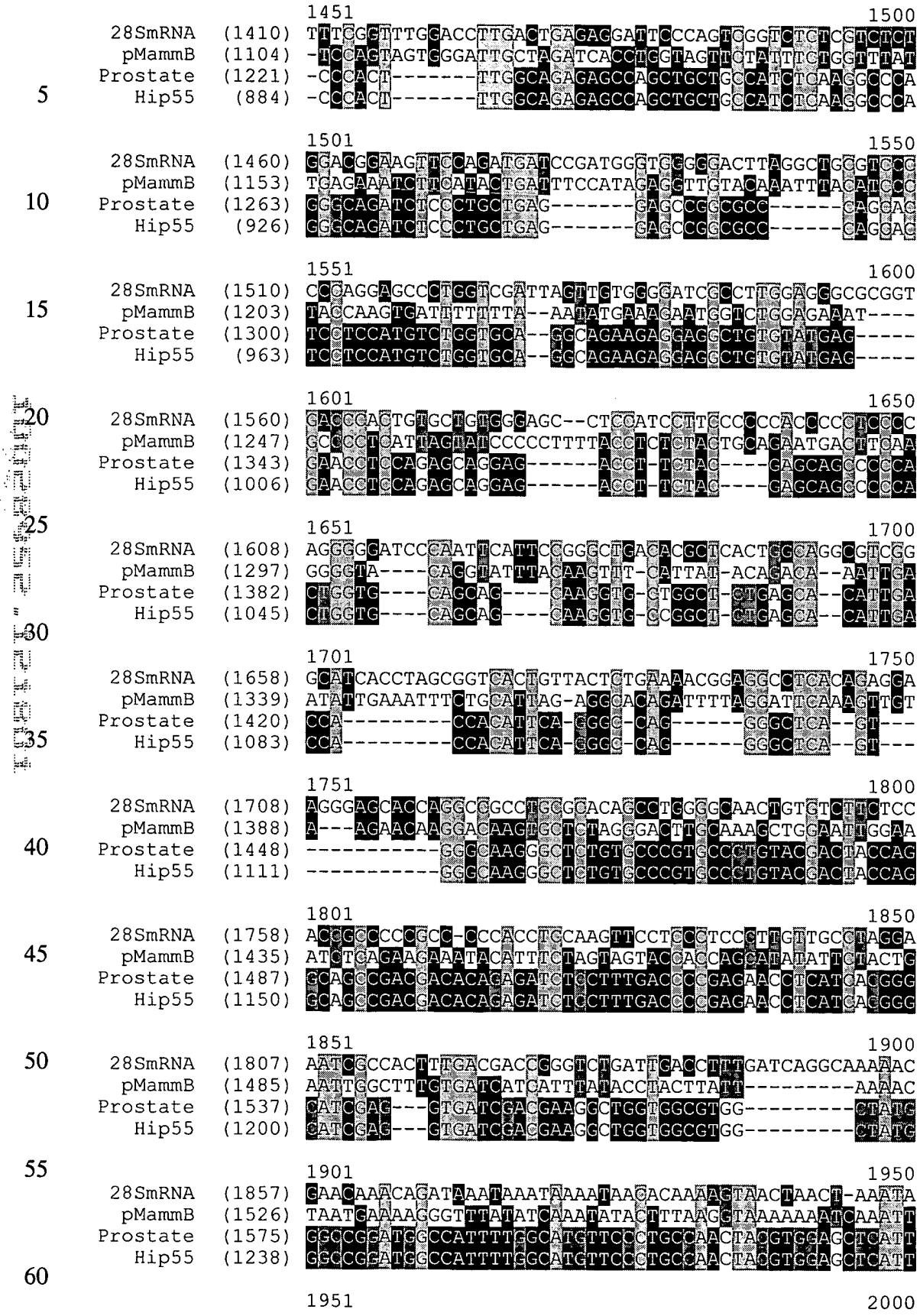
Table 2

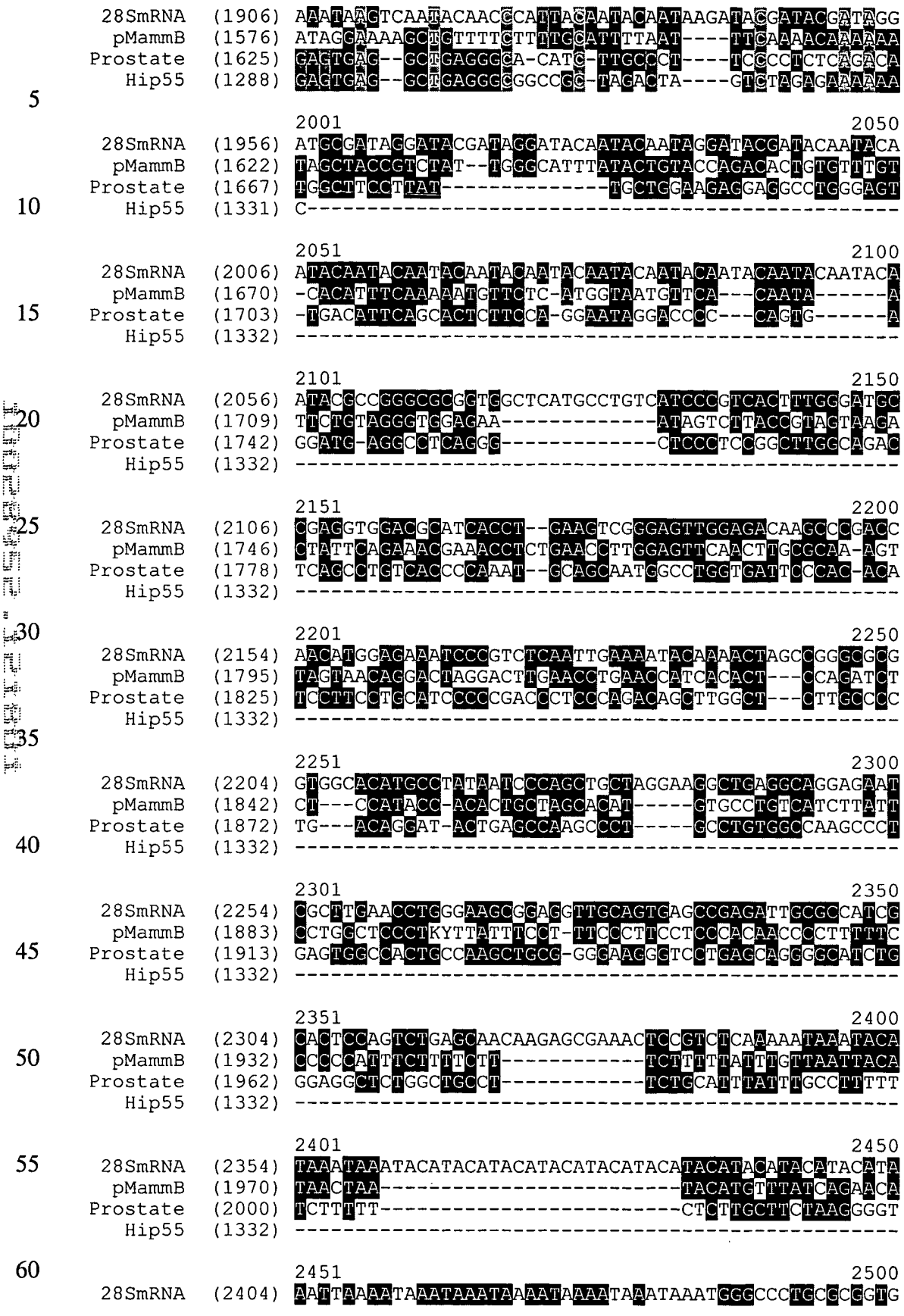
Prostate ECGI Homology




Hip55 (1) ATGGCGGCGAACCT GAGCCGGAACGGGCCAGCGC
451 500
28SmRNA (449) A-CTGGCGACCCGAGTGGATCTGCCCGCCCCGGCCTCCCAAAGTGCTGGG
pMammB (185) TTCTGGCG--CCAAGCG CCCGGCCGCGCGCCG-GCCGGG
Prostate (295) TTCTGGCG--CCAAGCG CCCGGCCGCGCGCCG-GCCGGG
Hip55 (35) TGCAAGAG--GCCTACG TGCGGGTGGTCACCGAGAAGTC
501 550
28SmRNA (498) -GTGACAGGCGTGAGCCATCGTGACTGGCCCGGCTACGTTTATTTATTTAT
pMammB (221) CGCGACCCGCTCCGGGGACAGTGCC--AGTGGGGAGTTTGACTGGGG---
Prostate (331) CGCGACCCGCTCCGGGGACAGTGCC-AGGTGGGGAGTTTGACTGGGG---
Hip55 (72) CCCGACCGACTGGCTCTCTTTACCTATGAAGGCAACAGCAATGACAT--
551 600
28SmRNA (547) TTTTTTTAATTATTTTACTTTTTTTTAGTTTTCCATTTTAATCTATTTATT
pMammB (266) CGGTACACCTGTCAAACGGTAACGCAGGTGTCC---TAAGGCGAGCTCAG
Prostate (377) CGGTACACCTGTCAAACGGTAACGCAGGTGTCC---TAAGGCGAGCTCAG
Hip55 (120) CCGCGTGGCTGGCACAGGGGAG----GGTGGCC---TGGAG-GAGATGGT
601 650
28SmRNA (597) TATTTACATTTATTTATTTATTTATTTATTTACTTATTTATTTATTTTCG
pMammB (313) GGAGGACA-AAACCTCCCGTGGAGCAGAAGGGCAAAA-------TGATCT
Prostate (424) GGAGGACAGAAACCTCCCGTGGAGCAGAAGGGCAAAAGCTCGCTTGATCT
Hip55 (162) GGAGGAGCTCAAC---------AGCGGGAAGG----------TGATGT
651 700
28SmRNA (647) AGACAGACTCTCGCTCTGCTGCCCAGGCTGGAGTGCAGCGGCGTGATC--
pMammB (355) TGATTTTCAGTACGAATACAGACCGTGAAAGCGGG--GCCTCA-GATC-T
Prostate (474) TGATTTTCAGTACGAATACAGACCGTGAAAGCGGG--GCCTCACGATCCT
Hip55 (191) ACGCCTTCTGCA-GAGTGAAGGACCCCAACTCTGG--ACTGCCCAAA---
701 750
28SmRNA (695) TCGGCTCAC-TGCAACGTCCGCCTCCCGGGTTCACGCCATTCTCCTGCCT
pMammB (401) TCTGACCTTTTGGGTTTTA-AGCAGGAGGTGTCAGAAAAGT-----TACCA
Prostate (522) TCTGACCTTTTGGGTTTTA-AGCAGGAGGTGTCAGAAAAGT-----TACCA
Hip55 (235) TTTGTCCTCATCAACTGGACAGGCGAGGGCGTGAACGATGT----GCGGA
751 800
28SmRNA (744) CAGCCTCCCAAGTAGCTGGGACTACAGGCGCCGGCCACCGTGCCCGGCTA
pMammB (446) CAGGGAT--AACTGGCTTGT------GGCGGCCA--AGCGTTCAAAGCGA
Prostate (567) CAGGGAT--AACTGGCTTGT------GGCGGCCA--AGCGTTCATAGCGA
Hip55 (281) -AGGGA--------GCCTGT------GCCAGCCA--CG---TCA--GGAC
801 850
28SmRNA (794) ACTTTTTGTATTTTGAGTAGAGATGGGGTTTCACTGTGGTAGCCAGGATG
pMammB (486) CGTCGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCATTGGGAAG---
Prostate (607) CGTCGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCATTGTGAAG---
Hip55 (309) CATGGCCAGCT--TCCT--GAAGGGGGCCCATGTGACCATCA--ACG---
851 900
28SmRNA (844) GTCTCGATCTCCTGACCCCGTGATCCGTCCACCTCGGCCTCCCAAA---G
pMammB (533) --CA-GAATTCACCAAGCGTTGGATTGTTCACCCACTAATAGGGAACGTG
Prostate (654) --CA-GAATTCACCAAGCGTTGGATTGTTCACCCACTAATAGGGAACGTG
Hip55 (350) --CACGGGCCGAGCAGGATGTGGAGCCTGAGTGCA-TCATGGAGAAGGTG
901 950
28SmRNA (891) TGCTGGCATG--ACAGGCGTGAGCCACC--GGCCCCGGCCTA---TTTAT
pMammB (580) AGCTGGGTTTAGACCGTCGTGAGACAGGTTTGT-TTACCCTACTGATGAT
Prostate (701) AGCTGGCATTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACTGATGAT
Hip55 (397) GCGAAGGCTT-------CAGGTGCCAACTACAGCTTTCAGAA--GGAGAG


951 1000
28SmRNA (934) CTATTTATTAACTTTGAGTCCAGGT---TATG-AAACCAGT--TAGTTTT
pMammB (629) GTGTTGTTGCCATGGTAATCCTGCTCAGTACG-AGAGGAACCGCAGGTTC
Prostate (751) GTGTTGTTGCCATGGTAATCCTGCTCAGTACG-AGAGGAACCGCAGGTTC
Hip55 (438) TGGCCGCTTCCAGGACGTGGGACCCAGGCCCAGTGGGCTCTGTGTACC
1001 1050
28SmRNA (978) TGTAATTTTTTTTTTTTTTTTTTTTTTTGAGACGAGGTTTCACCGTGTT
pMammB (678) AGACATTTGGTGTATGTG-CTTGGCTGGGGAGCCAATGG--GGCGAAGCT
Prostate (800) AGACATTTGGTGTATGTG-CTTGGCTGAGGAGCCAATGG--GGCGAAGCT
Hip55 (488) AGAA----GACCAATGC--CGTGTCTGAGATTAAAAGGGTTGGTAAAGAC
1051 1100
28SmRNA (1028) GCCAAGGCTTGGAC--CGAGGGATCCACCGGCCCTCGGCCTGCCAAAAGT
pMammB (725) ACCATCTGTGGGATTATTACTGAACGCCTCTAAGTCAGAATGCCGCCCAG
Prostate (847) ACCATCTGTGGGATTATGACTGAACGCCTCTAAGTCAGAATGCCGCCCAG
Hip55 (532) AGCTTCTGGGCCAA-AGCAGAGAAGGAGG--AGG--AGAACCGTCGGCTG
1101 1150
28SmRNA (1076) GCGGGGATGACAGGCGCGAGCGTACCGCGCC-CGGA--CCCCCCTTTCC
pMammB (775) GCGGA-ACGATACGGCAGCGCCG-CGGAGCCTCGGTTGGCCTCGGATGGC
Prostate (897) GCGGA-ACGATACGGCAGCGCCG-CGGAGCCTCGGTTGGCCTCGGATAGC
Hip55 (577) GAGGA-A--AAGCGGCGG-GCCG-AGGAGGCACAGC-GGCAGCTGG-AGC
1151 1200
28SmRNA (1123) CCTTCCCCCGCTTGTCTTC-CCGACAGAC--AGTTTCACGGCAGAGGGTT
pMammB (823) CGGTCGCCCCGCCTGTCCCCGCCGCGGGGGCGCCCCCCCCCTCCACGCGCC
Prostate (945) CGGTCGCCCCGCCTGTCCCCGCCGCGGGGGC-CGCCCCCCCCTCCACGCGCC
Hip55 (620) AGGAGCGCCCGGGAGCGTGAGCTG-CGTGA-GGCTGCACGCCGGCAGGAGC
1201 1250
28SmRNA (1170) TGGCTGGCGTGCTTAAACTGATTCTAAATAGAAATTTGGGAG----GTCA
pMammB (873) CCGCGCGCGCGGGAGGGCGCGTGCCCCGCCGCGCGCCGGGACCGGGGTCC
Prostate (994) CCGCGCGCGCGGGAGGGCGCGTGCCCCGCCGCGCGCCGGGACCGGGGTCC
Hip55 (668) GCTATCAGGAGCAGGGTGGCGAGGCCAGCCCCCAGA--GGACGTGGGAGC
1251 1300
28SmRNA (1216) GCTTCTG---GCCTCACGGACTCTGAGCCGAGGAGTCCCCTG---GTCTG
pMammB (923) GGTGCGCAGTGCCCTTCGTCCTGGGAAACGGGGCGCGGCCGGAAAGGCGG
Prostate (1044) GGTGCGCAGTGCCCTTCGTCCTGGGAAACGGGGCGCGGCCGGAAAGGCGG
Hip55 (716) AG--CAGCAAGAAGTGGTTTCAAGGAAGCCGAAATGAG-CAGGA--GTCTG
1301 1350
28SmRNA (1260) TCTATCAGAGGACCGTACACGTAAGGAGGAGAAAAATCGTAACGTTCAAA
pMammB (973) CCGCCCCCTCGCCCGT-CACGCACCG-CACGTTCGTGCT---CGTGCCGA
Prostate (1094) CCGCCCCCTCGCCCGT-CACGCACCG-CACGTTCGTGGGAACCTGGCGC
Hip55 (761) CCGTGCACCCGAGGGA-GATTTTCAA-GCAGAAGGAGAGGCCATGTCC-
1351 1400
28SmRNA (1310) GTCAGTCATTTTGTGATACAGAAATACACGGATTCACCCAAAACACAGAA
pMammB (1018) ATTCGGCACGAGTAGCACCATTCACAATAGA---CATACAAGTGCATGTA
Prostate (1142) -TAAACCACCTCCATCTCCAGTCGTCA---GC---CTGGCAAGCTGAGGAG
Hip55 (808) ----ACCACCTCCATCTCCAGTCGTCA---GC---CTGGCAAGCTGAGGAG
1401 1450
28SmRNA (1360) ACCAGTCTTTTAGAAATGGCCTTAGCCCTGGTGTCCGTGCCAGTGATTCT
pMammB (1065) TCT-TTATTATATAA------TGAATTCTTTT--CCTTTGGGGAGATA---
Prostate (1186) CCCCTTCCTGCAGAA------GCAGCTCAC----CCAACC---AGAGA---
Hip55 (849) CCCCTTCCTGCAGAA------GCAGCTCAC----CCAACC---AGAGA--


1451 1500
28SmRNA (1410) TTTCGGTTTGGACCTTGACTGAGAGGATTCCCAGTCGGTCTGTCGTCTCT
pMammB (1104) -TCCAGTAGTGGGATTGCTAGATCACCTGGTAGTTCTATTTGTGGTTTAT
Prostate (1221) -CCCACT-------TTGGCAGAGAGCCAGCTGCTGCCATCTCAAGGCCCA
Hip55 (884) -CCCACT-------TTGGCAGAGAGCCAGCTGCTGCCATCTCAAGGCCCA
1501 1550
28SmRNA (1460) GGACGGAAGTTCCAGATGATCCGATGGGTGGGGGACTTAGGCTGCGTCCG
pMammB (1153) TGAGAAATCTTCATACTGATTTCCATAGAGCTTGTACAAATTTACATCCC
Prostate (1263) GGGCAGATCTCCCTGCTGAG-------GAGCCGGCGCC------CAGGAG
Hip55 (926) GGGCAGATCTCCCTGCTGAG-------GAGCCGGCGCC------CAGGAG
1551 1600
28SmRNA (1510) CCCAGGAGCCCTGGTCGATTAGTTGTGGGGATCGCCTTGGAGCGGCGCGGT
pMammB (1203) TACCAAGTGATTTTTTTA--AATATGAAAGAATGGTCTGGAGAAAT----
Prostate (1300) TCCTCCATGTCTGGTGCA--GGCAGAAGAGGAGGCTGTGTATGAG-----
Hip55 (963) TCCTCCATGTCTGGTGCA--GGCAGAAGAGGAGGCTGTGTATGAG-----
1601 1650
28SmRNA (1560) GACCCACTGTGCTGTGGGAGC--CTCCATCCTTCCCCCCACCCCCTCCCC
pMammB (1247) GCCCCTCATTAGTATCCCCCTTTTACCTCTCTACTGCAGAATGACTTCAA
Prostate (1343) GAACCTCCAGAGCAGGAG------ACCT-TCTAC----GAGCAGCCCCCA
Hip55 (1006) GAACCTCCAGAGCAGGAG------ACCT-TCTAC----GAGCAGCCCCCA
1651 1700
28SmRNA (1608) AGGGGCATCCCAATTCATTCCGGGCTGACACGCTCACTGGCAGGCGTCGG
pMammB (1297) GGGGTA----CAGGTATTTACAAGTTT-CATTAT-ACAGACA--AATTCA
Prostate (1382) CTGGTG----CAGCAG----CAAGGTG-GTGGCT-GTGAGCA--CATTGA
Hip55 (1045) CTGGTG----CAGCAG----CAAGGTG-GCGGCT-GTGAGCA--CATTGA
1701 1750
28SmRNA (1658) GCATCACCTAGCGGTCACTGTTACTCTGAAAACGGAGGCCTCACAGAGGA
pMammB (1339) ATATTGAAATTTCTGCATTAG-AGGCACAGATTTTAGGATTCAAAGTTGT
Prostate (1420) CCA---------CCACATTCA-GGGC-CAG------GGGCTCA--GT---
Hip55 (1083) CCA---------CCACATTCA-GGGC-CAG------GGGCTCA--GT---
1751 1800
28SmRNA (1708) AGGGAGCACCAGGCCGCCTGCGCACAGCCTGGGGCAACTGTGTCTTCTCC
pMammB (1388) A---AGAACAAGGACAAGTGCTCTAGGGACTTGCAAAGCTGGAATTGGAA
Prostate (1448) ----------GGCCAAGGGCTCTGTGCCCGTGCCCTGTACGACTACCAG
Hip55 (1111) ----------GGCCAAGGGCTCTGTGCCCGTGCCCTGTACGACTACCAG
1801 1850
28SmRNA (1758) ACGCCCCCGCC-CCCACCTGCAAGTTCCTCCTCCGTTGTTGCCTAGGA
pMammB (1435) ATCTCAGAAGAAATACATTTCTAGTAGTACCACCAGCATATATTCTACTG
Prostate (1487) GCAGCCGACGACACAGAGATCTCCTTTGACCCCGAGAACCTCATCACGGG
Hip55 (1150) GCAGCCGACGACACAGAGATCTCCTTTGACCCCGAGAACCTCATCACGGG
1851 1900
28SmRNA (1807) AATCGCCACTTTGACGACCGGGTCTGATTCACCTTTGATCAGGCAAAAAC
pMammB (1485) AATTGGCTTTGTGATCATCATTTATACCTACTTATT---------AAAAC
Prostate (1537) CATCGAG---GTGATCGACGAAGGCTGGTGGCGTGG---------CTATG
Hip55 (1200) CATCGAG---GTGATCGACGAAGGCTGGTGGCGTGG---------CTATG
1901 1950
28SmRNA (1857) GAACAAACAGATAAATAAATAAAATAACACAAAAGTAACTAACT-AAATA
pMammB (1526) TAATGAAAAGGGTTTATATCAAATATACTTTAAGGTAAAAAATCAAATT
Prostate (1575) GGCCGGATGGCCATTTTGGCATGTTCCCTGCCAACTACGTGGAGCTCATT
Hip55 (1238) GGCCGGATGGCCATTTTGGCATGTTCCCTGCCAACTACGTGGAGCTCATT
1951 2000


28SmRNA (1906) AAATAAGTCAATACAACGCATTACAATACAATAAGATACGATACGATAGG
pMammB (1576) ATAGGAAAAGCTGTTTTCTTTTGCATTTTAAT----TTCAAAACAAAAA
Prostate (1625) GAGTGAG--GCTGAGGGCA-CATG-TTGCCCT----TCCCCTCTCAGACA
Hip55 (1288) GAGTGAG--GCTGAGGGCGGCCGG-TAGACTA----GTCTAGAGAAAAA
2001 2050
28SmRNA (1956) ATGCGATAGGATACGATAGGATACAATACAATAGGATACGATACAATACA
pMammB (1622) TAGCTACCGTCTAT--TGGGCATTTATACTGTACCAGACACTGTGTTTGT
Prostate (1667) TGGCTTCCTTAT----------TGCTGGAAGAGGAGGCCTGGGAGT
Hip55 (1331) C----------
2051 2100
28SmRNA (2006) ATACAATACAATACAATACAATACAATACAATACAATACAATACAATACA
pMammB (1670) -CACATTTCAAAAATGTTCTC-ATGGTAATGTTCA---CAATA------A
Prostate (1703) -TGACATTCAGCACTCTTCCA-GGAATAGGACCCC---CAGTG------A
Hip55 (1332) ----------
2101 2150
28SmRNA (2056) ATACGCCGGGCGCGGTGGCTCATGCCTGTCATCCCGTCACTTTGGGATGC
pMammB (1709) TTCTGTAGGGTGGAGAA----------ATAGTCTTACCGTAGTAAGA
Prostate (1742) GGATG-AGGCCTCAGGG----------CTCCCTCCGGCTTGGCAGAC
Hip55 (1332) ----------
2151 2200
28SmRNA (2106) CGAGGTGGACGCATCACCT--GAAGTCGGGAGTTGGAGACAAGCCCGACC
pMammB (1746) CTATTCAGAAACGAAACCTCTGAACCTTGGAGTTCAACTTGCGCAA-AGT
Prostate (1778) TCAGCCTGTCACCCCAAAT--GCAGCAATGGCCTGGTGATTCCCAC-ACA
Hip55 (1332) ----------
2201 2250
28SmRNA (2154) AACATGGAGAAATCCCGTCTCAATTGAAAATACAAAACTAGCCGGGCGCG
pMammB (1795) TAGTAACAGGACTAGGACTTGAACCTGAACCATCACACT---CCAGATCT
Prostate (1825) TCCTTCCTGCATCCCCGACCCTCCCAGACAGCTTGGCT---CTTGCCCC
Hip55 (1332) ----------
2251 2300
28SmRNA (2204) GTGGCACATGCCTATAATCCCAGCTGCTAGGAAGGCTGAGGCAGGAGAAT
pMammB (1842) CT---CCATACC-ACACTGCTAGCACAT-----GTGCCTGTCATCTTATT
Prostate (1872) TG---ACAGGAT-ACTGAGCCAAGCCCT-----GCCTGTGGCCAAGCCCT
Hip55 (1332) ----------
2301 2350
28SmRNA (2254) CGCTTGAACCTGGGAAGCGGAGGTTGCAGTGAGCCGAGATTGCGCCATCG
pMammB (1883) CCTGGCTCCCTKYTTATTTCCT-TTCCCTTCCTCCCACAACCCCTTTTTC
Prostate (1913) GAGTGGCCCACTGCCCAAGCTGCG-GGGAAGGGTCCTGAGCAGGGGCATCTG
Hip55 (1332) ----------
2351 2400
28SmRNA (2304) CACTCCAGTCTGAGCAACAAGAGCGAAACTCCGTCTCAAAAATAAATACA
pMammB (1932) CCCCCATTTCTTTTCTT----------TCTTTTTATTTGTTAATTACA
Prostate (1962) GGAGGCTCTGGCTGCCT----------TCTGCATTTATTTGCCTTTTT
Hip55 (1332) ----------
2401 2450
28SmRNA (2354) TAAATAAATACATACATACATACATACATACATACATACATACATACATA
pMammB (1970) TAACTAA----------TACATGTTTATCAGAACA
Prostate (2000) TCTTTTT----------CTCTTGCTTCTAAGGGT
Hip55 (1332) ----------
2451 2500
28SmRNA (2404) AATTAAAATAAATAAATAAATAAATAAATAAATGGGCCCTGCGCGGTG

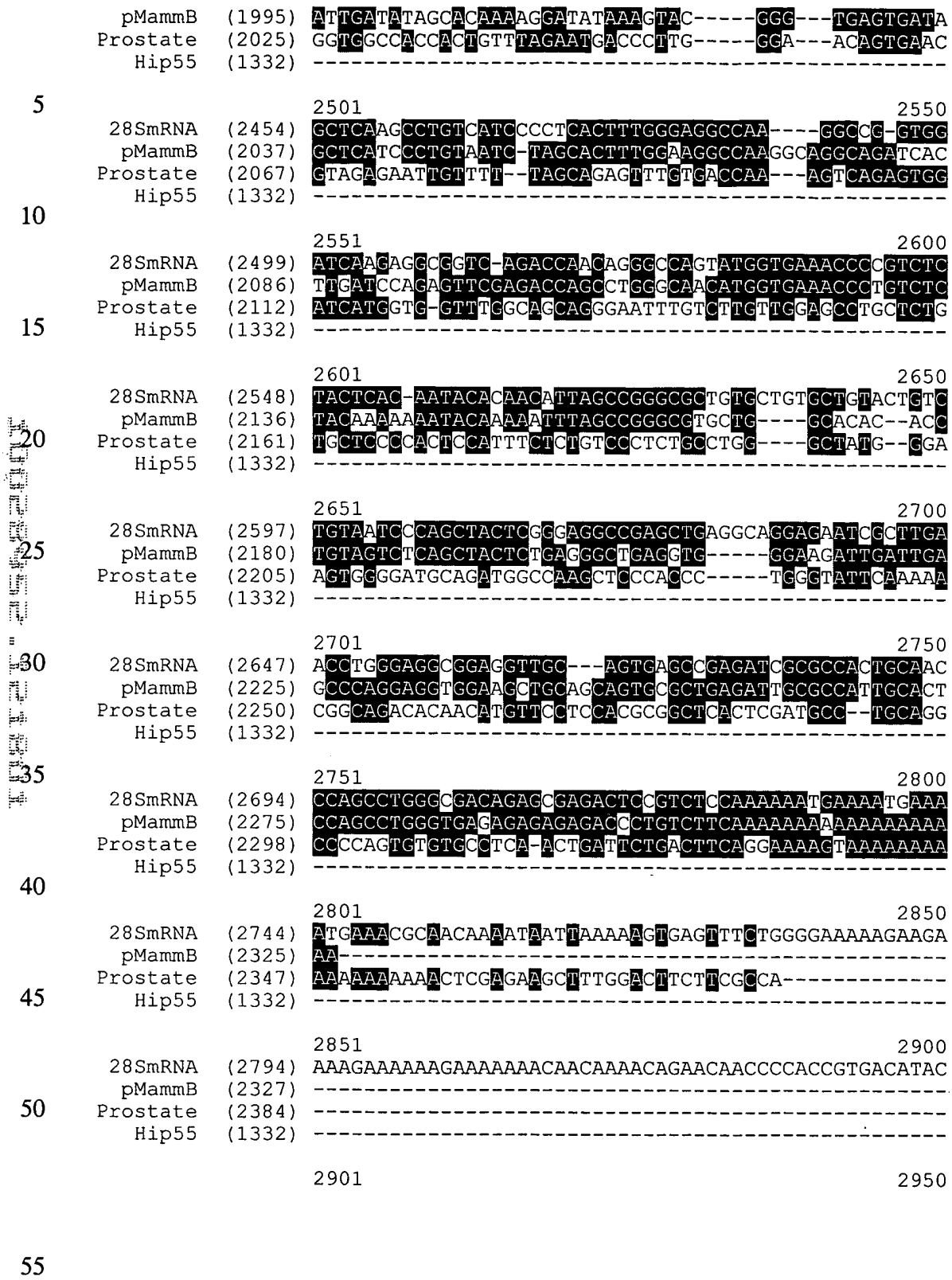

pMammB (1995) ATTGATATAGCACAAAAGGATATAAAGTAC-----GGG---TGAGTGATA
Prostate (2025) GGTGGCCACCACTGTTTAGAATGACCCTTG-----GGA---ACAGTGAAC
Hip55 (1332) ----------
2501 2550
28SmRNA (2454) GCTCAAGCCTGTCATCCCCTCACTTTGGGAGGCCAA----GGCCG-GTGG
pMammB (2037) GCTCATCCCTGTAATC-TAGCACTTTGGAAGGCCAAGGCAGGCAGATCAC
Prostate (2067) GTAGAGAATTGTTTT--TAGCAGAGTTTGTGACCAA---AGTCAGAGTGG
Hip55 (1332) ----------
2551 2600
28SmRNA (2499) ATCAAGAGGCGGTC-AGACCAACAGGGCCAGTATGGTGAAACCCCGTCTC
pMammB (2086) TTGATCCAGAGTTCGAGACCAGCCTGGGCAACATGGTGAAACCCTGTCTC
Prostate (2112) ATCATGGTG-GTTTGGCAGCAGGGAATTTGTCTTGTTGGAGCCTGCTCTG
Hip55 (1332) ----------
2601 2650
28SmRNA (2548) TACTCAC-AATACACAACATTAGCCGGGCGCTGTGCTGTGCTGTACTGTC
pMammB (2136) TACAAAAAAATACAAAATTTAGCCGGGCGTGCTG----GCACAC--ACC
Prostate (2161) TGCTCCCCACTCCATTTCTCTGTCCTCTGCCTGG----GCTATG--GGA
Hip55 (1332) ----------
2651 2700
28SmRNA (2597) TGTAATCCCAGCTACTCGGGAGGCCGAGCTGAGGCAGGAGAATCGCTTGA
pMammB (2180) TGTAGTCTCAGCTACTCTGAGGGCTGAGGTG-----GGAAGATTGATTGA
Prostate (2205) AGTGGGGATGCAGATGGCCAAGCTCCCACCC-----TGGGTATTCAAAAA
Hip55 (1332) ----------
2701 2750
28SmRNA (2647) ACCTGGGAGGCGGAGGTTGC---AGTGAGCCGAGATCGCGCCACTGCAAC
pMammB (2225) GCCCAGGAGGTGGAAGCTGCAGCAGTGCGCTGAGATTGCGCCATTGCACT
Prostate (2250) CGGCAGACACAACATGTTCCTCCACGCGGCTCACTCGATGCC--TGCAGG
Hip55 (1332) ----------
2751 2800
28SmRNA (2694) CCAGCCTGGGCGACAGAGCGAGACTCCGTCTCCAAAAAATGAAAATGAAA
pMammB (2275) CCAGCCTGGGTGAGAGAGAGAGACCCTGTCTTCAAAAAAAAAAAAAAAAA
Prostate (2298) CCCCAGTGTGTGCCTCA-ACTGATTCTGACTTCAGGAAAAGTAAAAAAAA
Hip55 (1332) ----------
2801 2850
28SmRNA (2744) ATGAAACGCAACAAAATAATTAAAAAGTGAGTTTCTGGGGAAAAAGAAGA
pMammB (2325) AA----------
Prostate (2347) AAAAAAAAAACTCGAGAAGCTTTGGACTTCTTCGCCA----------
Hip55 (1332) ----------
2851 2900
28SmRNA (2794) AAAGAAAAAAGAAAAAAACAACAAAACAGAACAACCCCACCGTGACATAC
pMammB (2327) ----------
Prostate (2384) ----------
Hip55 (1332) ----------
2901 2950

Table 3

Putative Prostate ECGI Amino Acid Sequence

```
          H  E  I   P  T  V  P   T  Y  Y   P  A  K   P  Q
     1    GCACGAGATT CCCACTGTCC CTACCTACTA TCCAGCGAAA CCACAGCCAA
          CGTGCTCTAA GGGTGACAGG GATGGATGAT AGGTCGCTTT GGTGTCGGTT
          ·  E  R  A   W  R  N   Q  R  G  K   K  T  L   L  S
    51    GGGAACGGGC TTGGCGGAAT CAGCGGGGAA AGAAGACCCT GTTGAGCTTG
          CCCTTGCCCG AACCGCCTTA GTCGCCCCTT TCTTCTGGGA CAACTCGAAC
           T  L  V  W   H  G  E   E  T  *   E  V  *  N   K  W
   101    ACTCTAGTCT GGCACGGTGA AGAGACATGA GAGGTGTAGA ATAAGTGGGA
          TGAGATCAGA CCGTGCCACT TCTCTGTACT CTCCACATCT TATTCACCCT
          ·  A  P  G   A  P  P  V   S  P  R   G  A  R   G  G
   151    GGCCCCCGGC GCCCCCCCGG TGTCCCCGCG AGGGGCCCGG GGCGGGGTCC
          CCGGGGGCCG CGGGGGGGCC ACAGGGGCGC TCCCCGGGCC CCGCCCCAGG
          ·  R  P  C   G  P  P   V  K  Y  H   Y  S  D   R  F
   201    GCCGGCCCTG CGGGCCGCCG GTGAAATACC ACTACTCTGA TCGTTTTTTC
          CGGCCGGGAC GCCCGGCGGC CACTTTATGG TGATGAGACT AGCAAAAAAG
           T  D  P  V   R  R  G   G  E  P   R  G  A  L   A  S
   251    ACTGACCCGG TGAGGCGGGG GGGCGAGCCC CGAGGGGCTC TCGCTTCTGG
          TGACTGGGCC ACTCCGCCCC CCCGCTCGGG GCTCCCCGAG AGCGAAGACC
          ·  A  K  R   P  A  A  R   R  P  G   A  T  R   S  G
   301    CGCCAAGCGC CCGGCCGCGC GCCGGCCGGG CGCGACCCGC TCCGGGGACA
          GCGGTTCGCG GGCCGGCGCG CGGCCGGCCC GCGCTGGGCG AGGCCCCTGT
          ·  A  R  W   G  V  *   L  G  R  Y   T  C  Q   T  V
   351    GTGCCAGGTG GGGAGTTTGA CTGGGGCGGT ACACCTGTCA AACGGTAACG
          CACGGTCCAC CCCTCAAACT GACCCCGCCA TGTGGACAGT TTGCCATTGC
           Q  V  S  *   G  E  L   R  E  D   R  N  L  P   W  S
   401    CAGGTGTCCT AAGGCGAGCT CAGGGAGGAC AGAAACCTCC CGTGGAGCAG
          GTCCACAGGA TTCCGCTCGA GTCCCTCCTG TCTTTGGAGG GCACCTCGTC
          ·  R  A  K   A  R  L  I   L  I  F   S  T  N   T  D
   451    AAGGGCAAAA GCTCGCTTGA TCTTGATTTT CAGTACGAAT ACAGACCGTG
          TTCCCGTTTT CGAGCGAACT AGAACTAAAA GTCATGCTTA TGTCTGGCAC
          ·  S  G  A   S  R  S   F  *  P  F   G  F  *   A  G
   501    AAAGCGGGGC CTCACGATCC TTCTGACCTT TTGGGTTTTA AGCAGGAGGT
          TTTCGCCCCG GAGTGCTAGG AAGACTGGAA AACCCAAAAT TCGTCCTCCA
           V  R  K  V   T  T  G   I  T  G   L  W  R  P   S  V
   551    GTCAGAAAAG TTACCACAGG GATAACTGGC TTGTGGCGGC CAAGCGTTCA
          CAGTCTTTTC AATGGTGTCC CTATTGACCG AACACCGCCG GTTCGCAAGT
          ·  S  D  V   A  F  *  S   F  D  V   G  S  S   Y  H
   601    TAGCGACGTC GCTTTTTGAT CCTTCGATGT CGGCTCTTCC TATCATTGTG
          ATCGCTGCAG CGAAAAACTA GGAAGCTACA GCCGAGAAGG ATAGTAACAC
          ·  A  E  F   T  K  R   W  I  V  H   P  L  I   G  N
   651    AAGCAGAATT CACCAAGCGT TGGATTGTTC ACCCACTAAT AGGGAACGTG
          TTCGTCTTAA GTGGTTCGCA ACCTAACAAG TGGGTGATTA TCCCTTGCAC
           S  W  D  *   T  V  V   R  Q  V   S  F  T  L   L  M
   701    AGCTGGGATT AGACCGTCGT GAGACAGGTT AGTTTTACCC TACTGATGAT
          TCGACCCTAA TCTGGCAGCA CTCTGTCCAA TCAAAATGGG ATGACTACTA
          ·  C  C  C   H  G  N  P   A  Q  Y   E  R  N   R  R
   751    GTGTTGTTGC CATGGTAATC CTGCTCAGTA CGAGAGGAAC CGCAGGTTCA
          CACAACAACG GTACCATTAG GACGAGTCAT GCTCTCCTTG GCGTCCAAGT
          ·  H  L  V   Y  V  L   G  *  G  A   N  G  A   K  L
   801    GACATTTGGT GTATGTGCTT GGCTGAGGAG CCAATGGGGC GAAGCTACCA
```

```
        CTGTAAACCA CATACACGAA CCGACTCCTC GGTTACCCCG CTTCGATGGT
         S  V  G  L  *  L  N  A  S  K  S  E  S  R  P  G
  851   TCTGTGGGAT TATGACTGAA CGCCTCTAAG TCAGAATCCC GCCCAGGCGG
        AGACACCCTA ATACTGACTT GCGGAGATTC AGTCTTAGGG CGGGTCCGCC
        · T  I  R  Q  R  R  G  A  S  V  G  L  G  *  P
  901   AACGATACGG CAGCGCCGCG GAGCCTCGGT TGGCCTCGGA TAGCCGGTCC
        TTGCTATGCC GTCGCGGCGC CTCGGAGCCA ACCGGAGCCT ATCGGCCAGG
        · R  L  S  P  P  A  G  R  P  P  L  H  A  P  R
  951   CCCGCCTGTC CCCGCCGGCG GGCCGCCCCC CCCTCCACGC GCCCCGCGCG
        GGGCGGACAG GGGCGGCCGC CCGGCGGGGG GGGAGGTGCG CGGGGCGCGC
         R  G  R  A  R  A  P  P  R  A  G  T  G  V  R  C
 1001   CGCGGGAGGG CGCGTGCCCC GCCGCGCGCC GGGACCGGGG TCCGGTGCGG
        GCGCCCTCCC GCGCACGGGG CGGCGCGCGG CCCTGGCCCC AGGCCACGCC
        · V  P  F  V  L  G  N  G  A  R  P  E  R  R  P
 1051   AGTGCCCTTC GTCCTGGGAA ACGGGGCGCG GCCGGAAAGG CGGCCGCCCC
        TCACGGGAAG CAGGACCCTT TGCCCCGCGC CGGCCTTTCC GCCGGCGGGG
        · R  P  S  R  T  A  R  S  W  G  T  W  R  T
 1101   CTCGCCCGTC ACGCACCGCA CGTTCGTGGG GAACCTGGCG CTAAACCACC
        GAGCGGGCAG TGCGTGGCGT GCAAGCACCC CTTGGACCGC GATTTGGTGG
         S  I  S  S  P  Q  P  G  K  L  R  S  P  F  L  Q
 1151   TCCATCTCCA GTCCTCAGCC TGGCAAGCTG AGGAGCCCCT TCCTGCAGAA
        AGGTAGAGGT CAGGAGTCGG ACCGTTCGAC TCCTCGGGGA AGGACGTCTT
        · Q  L  T  Q  P  E  T  H  F  G  R  E  P  A  A
 1201   GCAGCTCACC CAACCAGAGA CCCACTTTGG CAGAGAGCCA GCTGCTGCCA
        CGTCGAGTGG GTTGGTCTCT GGGTGAAACC GTCTCTCGGT CGACGACGGT
        · S  R  P  R  A  D  L  P  A  E  E  P  A  P  S
 1251   TCTCAAGGCC CAGGGCAGAT CTCCCTGCTG AGGAGCCGGC GCCCAGCACT
        AGAGTTCCGG GTCCCGTCTA GAGGGACGAC TCCTCGGCCG CGGGTCGTGA
         P  P  C  L  V  Q  A  E  E  E  A  V  Y  E  E  P
 1301   CCTCCATGTC TGGTGCAGGC AGAAGAGGAG GCTGTGTATG AGGAACCTCC
        GGAGGTACAG ACCACGTCCG TCTTCTCCTC CGACACATAC TCCTTGGAGG
        · E  Q  E  T  F  Y  E  Q  P  P  L  V  Q  Q  Q
 1351   AGAGCAGGAG ACCTTCTACG AGCAGCCCCC ACTGGTGCAG CAGCAAGGTG
        TCTCGTCCTC TGGAAGATGC TCGTCGGGGG TGACCACGTC GTCGTTCCAC
        · G  S  E  H  I  D  H  H  I  Q  G  Q  G  L  S
 1401   CTGGCTCTGA GCACATTGAC CACCACATTC AGGGCCAGGG GCTCAGTGGG
        GACCGAGACT CGTGTAACTG GTGGTGTAAG TCCCGGTCCC CGAGTCACCC
         Q  G  L  C  A  R  A  L  Y  D  Y  Q  A  A  D  D
 1451   CAAGGGCTCT GTGCCCGTGC CCTGTACGAC TACCAGGCAG CCGACGACAC
        GTTCCCGAGA CACGGGCACG GGACATGCTG ATGGTCCGTC GGCTGCTGTG
        · E  I  S  F  D  P  E  N  L  I  T  G  I  E  V
 1501   AGAGATCTCC TTTGACCCCG AGAACCTCAT CACGGGCATC GAGGTGATCG
        TCTCTAGAGG AAACTGGGGC TCTTGGAGTA GTGCCCGTAG CTCCACTAGC
        · E  G  W  W  R  G  Y  G  P  D  G  H  F  G  M
 1551   ACGAAGGCTG GTGGCGTGGC TATGGGCCGG ATGGCCATTT TGGCATGTTC
        TGCTTCCGAC CACCGCACCG ATACCCGGCC TACCGGTAAA ACCGTACAAG
         P  A  N  Y  V  E  L  I  E  *  G  *  G  H  I  L
 1601   CCTGCCAACT ACGTGGAGCT CATTGAGTGA GGCTGAGGGC ACATCTTGCC
        GGACGGTTGA TGCACCTCGA GTAACTCACT CCGACTCCCG TGTAGAACGG
        · F  P  S  Q  T  W  L  P  Y  C  W  K  R  R  P
 1651   CTTCCCCTCT CAGACATGGC TTCCTTATTG CTGGAAGAGG AGGCCTGGGA
        GAAGGGGAGA GTCTGTACCG AAGGAATAAC GACCTTCTCC TCCGGACCCT
        · *  H  S  A  L  F  Q  E  *  D  P  Q  *  G  *
 1701   GTTGACATTC AGCACTCTTC CAGGAATAGG ACCCCCAGTG AGGATGAGGC
        CAACTGTAAG TCGTGAGAAG GTCCTTATCC TGGGGGTCAC TCCTACTCCG
         L  R  A  P  S  G  L  A  D  S  A  C  H  P  K  C
 1751   CTCAGGGCTC CCTCCGGCTT GGCAGACTCA GCCTGTCACC CCAAATGCAG
        GAGTCCCGAG GGAGGCCGAA CCGTCTGAGT CGGACAGTGG GGTTTACGTC
        · N  G  L  V  I  P  T  H  P  S  C  I  P  R  P
 1801   CAATGGCCTG GTGATTCCCA CACATCCTTC CTGCATCCCC CGACCCTCCC
        GTTACCGGAC CACTAAGGGT GTGTAGGAAG GACGTAGGGG GCTGGGAGGG
```

```
        ·  T  A  W    L  L  P    L  T  G  Y    *  A  K    P  C
1851    AGACAGCTTG GCTCTTGCCC CTGACAGGAT ACTGAGCCAA GCCCTGCCTG
        TCTGTCGAAC CGAGAACGGG GACTGTCCTA TGACTCGGTT CGGGACGGAC
         W  P  S  P    E  W  P    L  P  S    C  G  E  G    S  *
1901    TGGCCAAGCC CTGAGTGGCC ACTGCCAAGC TGCGGGGAAG GGTCCTGAGC
        ACCGGTTCGG GACTCACCGG TGACGGTTCG ACGCCCCTTC CCAGGACTCG
        ·  G  A  S    G  R  L  W    L  P  S    A  F  I    C  L
1951    AGGGGCATCT GGGAGGCTCT GGCTGCCTTC TGCATTTATT TGCCTTTTTT
        TCCCCGTAGA CCCTCCGAGA CCGACGGAAG ACGTAAATAA ACGGAAAAAA
        ·  F  S  L    A  S  K    G  W  W  P    P  L  F    R  M
2001    CTTTTTCTCT TGCTTCTAAG GGGTGGTGGC CACCACTGTT TAGAATGACC
        GAAAAAGAGA ACGAAGATTC CCCACCACCG GTGGTGACAA ATCTTACTGG
         L  G  N  S    E  R  R    E  L  F    L  A  E  F    V  T
2051    CTTGGGAACA GTGAACGTAG AGAATTGTTT TTAGCAGAGT TTGTGACCAA
        GAACCCTTGT CACTTGCATC TCTTAACAAA AATCGTCTCA AACACTGGTT
        ·  V  R  V    D  H  G  G    L  A  A    G  N  L    S  C
2101    AGTCAGAGTG GATCATGGTG GTTTGGCAGC AGGGAATTTG TCTTGTTGGA
        TCAGTCTCAC CTAGTACCAC CAAACCGTCG TCCCTTAAAC AGAACAACCT
        ·  L  L  C    A  P  H    S  I  S  L    S  L  C    L  G
2151    GCCTGCTCTG TGCTCCCCAC TCCATTTCTC TGTCCCTCTG CCTGGGCTAT
        CGGACGAGAC ACGAGGGGTG AGGTAAAGAG ACAGGGAGAC GGACCCGATA
         G  K  W  G    C  R  W    P  S  S    H  P  G  Y    S  K
2201    GGGAAGTGGG GATGCAGATG GCCAAGCTCC CACCCTGGGT ATTCAAAAAC
        CCCTTCACCC CTACGTCTAC CGGTTCGAGG GTGGGACCCA TAAGTTTTTG
        ·  A  D  T    T  C  S  S    T  R  L    T  R  C    L  Q
2251    GGCAGACACA ACATGTTCCT CCACGCGGCT CACTCGATGC CTGCAGGCCC
        CCGTCTGTGT TGTACAAGGA GGTGCGCCGA GTGAGCTACG GACGTCCGGG
        ·  V  C  A    S  T  D    S  D  F  R    K  S  K    K  K
2301    CAGTGTGTGC CTCAACTGAT TCTGACTTCA GGAAAAGTAA AAAAAAAAAA
        GTCACACACG GAGTTGACTA AGACTGAAGT CCTTTTCATT TTTTTTTTTT
         K  K  L  E    K  L  W    T  S  S
2351    AAAAAACTCG AGAAGCTTTG GACTTCTTCG CCA
        TTTTTTGAGC TCTTCGAAAC CTGAAGAAGC GGT
```

Table 4

Putative MammC Amino Acid Sequence

```
          I  R  H   E  H  G   E  E  T  *   E  V  *   N  K
     1    GAATTCGGCA CGAGCACGGT GAAGAGACAT GAGAGGTGTA GAATAAGTGG
          CTTAAGCCGT GCTCGTGCCA CTTCTCTGTA CTCTCCACAT CTTATTCACC
           E  A  P  G   A  P  P   V  S  P   R  G  A  R   G  G
    51    GAGGCCCCCG GCGCCCCCCC GGTGTCCCCG CGAGGGGCCC GGGGCGGGGT
          CTCCGGGGGC CGCGGGGGGG CCACAGGGGC GCTCCCCGGG CCCCGCCCCA
          · R  R  P   C  G  P   P  V  K  Y   H  Y  S   D  R
   101    CCGCCGGCCC TGCGGGCCGC CGGTGAAATA CCACTACTCT GATCGTTTTT
          GGCGGCCGGG ACGCCCGGCG GCCACTTTAT GGTGATGAGA CTAGCAAAAA
          · T  D  P   V  R  R   G  G  E  P   R  G  A   L  A
   151    TCACTGACCC GGTGAGGCGG GGGGGCGAGC CCCGAGGGGC TCTCGCTTCT
          AGTGACTGGG CCACTCCGCC CCCCCGCTCG GGGCTCCCCG AGAGCGAAGA
           G  A  K  R   P  A  A   R  R  P   G  A  T  R   S  G
   201    GGCGCCAAGC GCCCGGCCGC GCGCCGGCCG GGCGCGACCC GCTCCGGGGA
          CCGCGGTTCG CGGGCCGGCG CGCGGCCGGC CCGCGCTGGG CGAGGCCCCT
          · S  A  R   W  G  V   *  L  G  R   Y  T  C   Q  T
   251    CAGTGCCAGG TGGGGAGTTT GACTGGGGCG GTACACCTGT CAAACGGTAA
          GTCACGGTCC ACCCCTCAAA CTGACCCCGC CATGTGGACA GTTTGCCATT
          · Q  V  S   *  G  E   L  R  E  D   R  N  L   P  W
   301    CGCAGGTGTC CTAAGGCGAG CTCAGGGAGG ACAGAAACCT CCCGTGGAGC
          GCGTCCACAG GATTCCGCTC GAGTCCCTCC TGTCTTTGGA GGGCACCTCG
           R  R  A  K   A  R  L   I  L  I   F  S  T  N   T  D
   351    AGAAGGGCAA AAGCTCGCTT GATCTTGATT TTCAGTACGA ATACAGACCG
          TCTTCCCGTT TTCGAGCGAA CTAGAACTAA AAGTCATGCT TATGTCTGGC
          · E  S  G   A  S  R   S  F  *  P   F  G  F   *  A
   401    TGAAAGCGGG GCCTCACGAT CCTTCTGACC TTTTGGGTTT TAAGCAGGAG
          ACTTTCGCCC CGGAGTGCTA GGAAGACTGG AAAACCCAAA ATTCGTCCTC
          · V  R  K   V  T  T   G  I  T  G   L  W  R   P  S
   451    GTGTCAGAAA AGTTACCACA GGGATAACTG GCTTGTGGCG GCCAAGCGTT
          CACAGTCTTT TCAATGGTGT CCCTATTGAC CGAACACCGC CGGTTCGCAA
           H  S  D  V   A  F  *   S  F  D   V  G  S  S   Y  H
   501    CATAGCGACG TCGCTTTTTG ATCCTTCGAT GTCGGCTCTT CCTATCATTG
          GTATCGCTGC AGCGAAAAAC TAGGAAGCTA CAGCCGAGAA GGATAGTAAC
          · E  A  E   F  T  K   R  W  I  V   H  P  L   I  G
   551    TGAAGCAGAA TTCACCAAGC GTTGGATTGT TCACCCACTA ATAGGGAACG
          ACTTCGTCTT AAGTGGTTCG CAACCTAACA AGTGGGTGAT TATCCCTTGC
          · S  W  V   *  T  V   V  R  Q  V   S  F  T   L  L
   601    TGAGCTGGGT TTAGACCGTC GTGAGACAGG TTAGTTTTAC CCTACTGATG
          ACTCGACCCA AATCTGGCAG CACTCTGTCC AATCAAAATG GGATGACTAC
           M  C  C  C   H  G  N   P  A  Q   Y  E  R  N   R  R
   651    ATGTGTTGTT GCCATGGTAA TCCTGCTCAG TACGAGAGGA ACCGCAGGTT
          TACACAACAA CGGTACCATT AGGACGAGTC ATGCTCTCCT TGGCGTCCAA
          · R  H  L   V  Y  V   L  G  *  G   A  N  G   A  K
   701    CAGACATTTG GTGTATGTGC TTGGCTGAGG AGCCAATGGG GCGAAGCTAC
          GTCTGTAAAC CACATACACG AACCGACTCC TCGGTTACCC CGCTTCGATG
          · S  V  G   L  *  L   N  A  S  K   S  E  S   R  P
   751    CATCTGTGGG ATTATGACTG AACGCCTCTA AGTCAGAATC CCGCCCAGGC
          GTAGACACCC TAATACTGAC TTGCGGAGAT TCAGTCTTAG GGCGGGTCCG
           G  T  I  R   Q  R  R   G  A  S   V  G  L  G   *  P
   801    GGAACGATAC GGCAGCGCCG CGGAGCCTCG GTTGGCCTCG GATAGCCGGT
          CCTTGCTATG CCGTCGCGGC GCCTCGGAGC CAACCGGAGC CTATCGGCCA
```

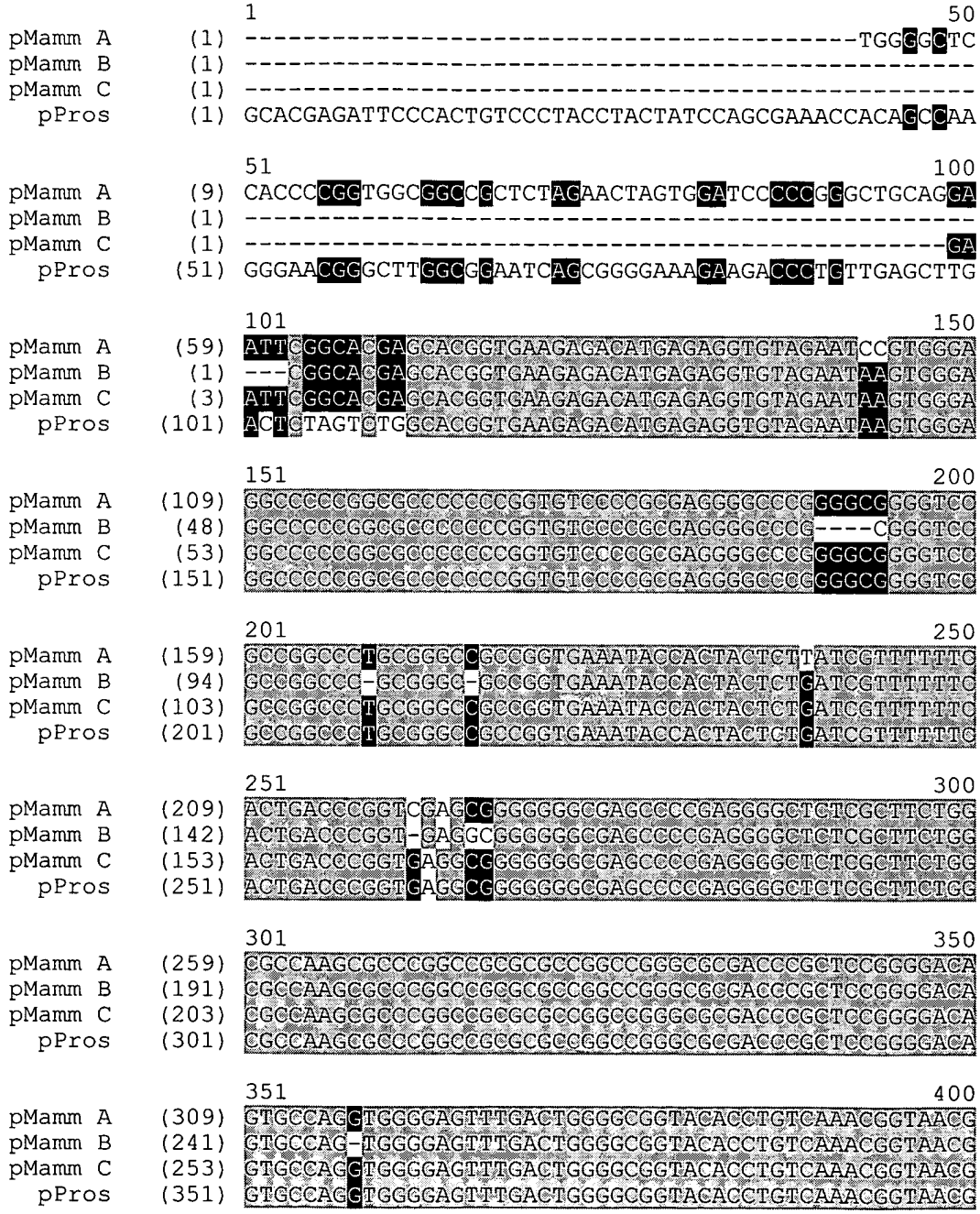
```
        · P  R  L   S  P  P  A   G  R  P   P  P  S   T  R
  851   CCCCCGCCTG TCCCCGCCGG CGGGCCGCCC CCCCCCCTCC ACGCGCCCCG
        GGGGGCGGAC AGGGGCGGCC GCCCGGCGGG GGGGGGGAGG TGCGCGGGGC
        · R  A  G   G  R  V   P  R  R  A   P  G  P   G  S
  901   CGCGCGCGGG AGGGCGCGTG CCCCGCCGCG CGCCGGGACC GGGGTCCGGT
        GCGCGCGCCC TCCCGCGCAC GGGGCGGCGC GCGGCCCTGG CCCCAGGCCA
         A  E  C  P   S  S  W   E  T  G   R  G  R  K   G  G
  951   GCGGAGTGCC CTTCGTCCTG GGAAACGGGG CGCGGCCGGA AAGGCGGCCG
        CGCCTCACGG GAAGCAGGAC CCTTTGCCCC GCGCCGGCCT TTCCGCCGGC
        · P  L  A   R  H  A  P   H  V  R   A  R  A   E  F
 1001   CCCCCTCGCC CGTCACGCAC CGCACGTTCG TGCTCGTGCC GAATTCGGCA
        GGGGGAGCGG GCAGTGCGTG GCGTGCAAGC ACGAGCACGG CTTAAGCCGT
        · S  S  T   I  H  N   R  H  T  S   A  C  I   F  M
 1051   CGAGTAGCAC CATTCACAAT AGACATACAA GTGCATGTAT CTTTATGATA
        GCTCATCGTG GTAAGTGTTA TCTGTATGTT CACGTACATA GAAATACTAT
         *  *  I  L   F  L  W   V  D  I   Q  *  W  D   C  *
 1101   TAATGAATTC TTTTCCTTTG GGTAGATATC CAGTAGTGGG ATTGCTAGAT
        ATTACTTAAG AAAAGGAAAC CCATCTATAG GTCATCACCC TAACGATCTA
        · T  W  *   F  Y  F  W   F  I  E   K  S  S   Y  *
 1151   CACCTGGTAG TTCTATTTCT GGTTTATTGA GAAATCTTCA TACTGATTTC
        GTGGACCATC AAGATAAAGA CCAAATAACT CTTTAGAAGT ATGACTAAAG
        · *  R  L   Y  K  F   T  S  L  P   S  D  F   F  K
 1201   CATAGAGGTT GTACAAATTT ACATCCCTAC CAAGTGATTT TTTTAAATAT
        GTATCTCCAA CATGTTTAAA TGTAGGGATG GTTCACTAAA AAAATTTATA
         E  R  M  V   W  R  N   A  P  H   *  Y  P  P   F  T
 1251   GAAAGAATGG TCTGGAGAAA TGCCCCTCAT TAGTATCCCC CTTTTACCTC
        CTTTCTTACC AGACCTCTTT ACGGGGAGTA ATCATAGGGG GAAAATGGAG
        · L  L  Q   N  D  F  K   G  Y  R   Y  L  Q   V  S
 1301   TCTACTGCAG AATGACTTCA AGGGGTACAG GTATTTACAA GTTTCATTAT
        AGATGACGTC TTACTGAAGT TCCCCATGTC CATAAATGTT CAAAGTAATA
        · R  Q  I   E  Y  *   N  F  C  I   R  G  T   D  F
 1351   ACAGACAAAT TGAATATTGA AATTTCTGCA TAAGAGGCAC AGATTTTAGG
        TGTCTGTTTA ACTTATAACT TTAAAGACGT ATTCTCCGTG TCTAAAATCC
         I  Q  S  C   M  N  K   D  K  C   S  R  D  L   Q  S
 1401   ATTCAAAGTT GTATGAACAA GGACAAGTGC TCTAGGGACT TGCAAAGCTG
        TAAGTTTCAA CATACTTGTT CCTGTTCACG AGATCCCTGA ACGTTTCGAC
        · N  W  K   S  Q  M  K   Y  I  S   S  S  T   T  S
 1451   GAATTGGAAA TCTCAGATGA AATACATTTC TAGTAGTACC ACCAGCATAT
        CTTAACCTTT AGAGTCTACT TTATGTAAAG ATCATCATGG TGGTCGTATA
        · S  T  E   L  A  L   *  S  S  L   I  P  T   Y  *
 1501   ATTCTACTGA ATTGGCTTTG TGATCATCAT TAATACCTAC TTATTAAAAC
        TAAGATGACT TAACCGAAAC ACTAGTAGTA ATTATGGATG AATAATTTTG
         *  *  K  G   F  I  S   N  I  L   *  G  I  K   I  K
 1551   TAATGAAAAG GGTTTATATC AAATATACTT TAAGGTATAA AAATCAAATT
        ATTACTTTTC CCAAATATAG TTTATATGAA ATTCCATATT TTTAGTTTAA
        · *  V  K   L  F  S  L   A  F  *   F  Q  N   I  K
 1601   ATAGGTAAAG CTGTTTTCTT TAGCATTTTA ATTTCAAAAC ATAAAATAGC
        TATCCATTTC GACAAAAGAA ATCGTAAAAT TAAAGTTTTG TATTTTATCG
        · P  S  I   G  H  L   Y  C  T  R   H  C  V   C  H
 1651   TACCGTCTAT TGGGCATTTA TACTGTACCA GACACTGTGT TTGTCACATT
        ATGGCAGATA ACCCGTAAAT ATGACATGGT CTGTGACACA AACAGTGTAA
         S  K  M  F   S  W  *   C  S  Q   *  F  C  R   V  R
 1701   TCAAAAATGT TCTCATGGTA ATGTTCACAA TAATTCTGTA GGGTGAGAAA
        AGTTTTTACA AGAGTACCAT TACAAGTGTT ATTAAGACAT CCCACTCTTT
        · S  L  T   V  V  R  L   F  S  K   R  N  L   *  T
 1751   TAGTCTTACC GTAGTAAGAC TATTCAGTAA ACGAAACCTC TGAACCTTGG
        ATCAGAATGG CATCATTCTG ATAAGTCATT TGCTTTGGAG ACTTGGAACC
        · F  N  L   R  K  V   S  N  R  T   R  T  *   T  *
 1801   AGTTCAACTT GCGCAAAGTT AGTAACAGGA CTAGGACTTG AACCTGAACC
        TCAAGTTGAA CGCGTTTCAA TCATTGTCCT GATCCTGAAC TTGGACTTGG
         I  T  L  Q   I  S  P   Y  H  T   A  S  T  C   A  C
```

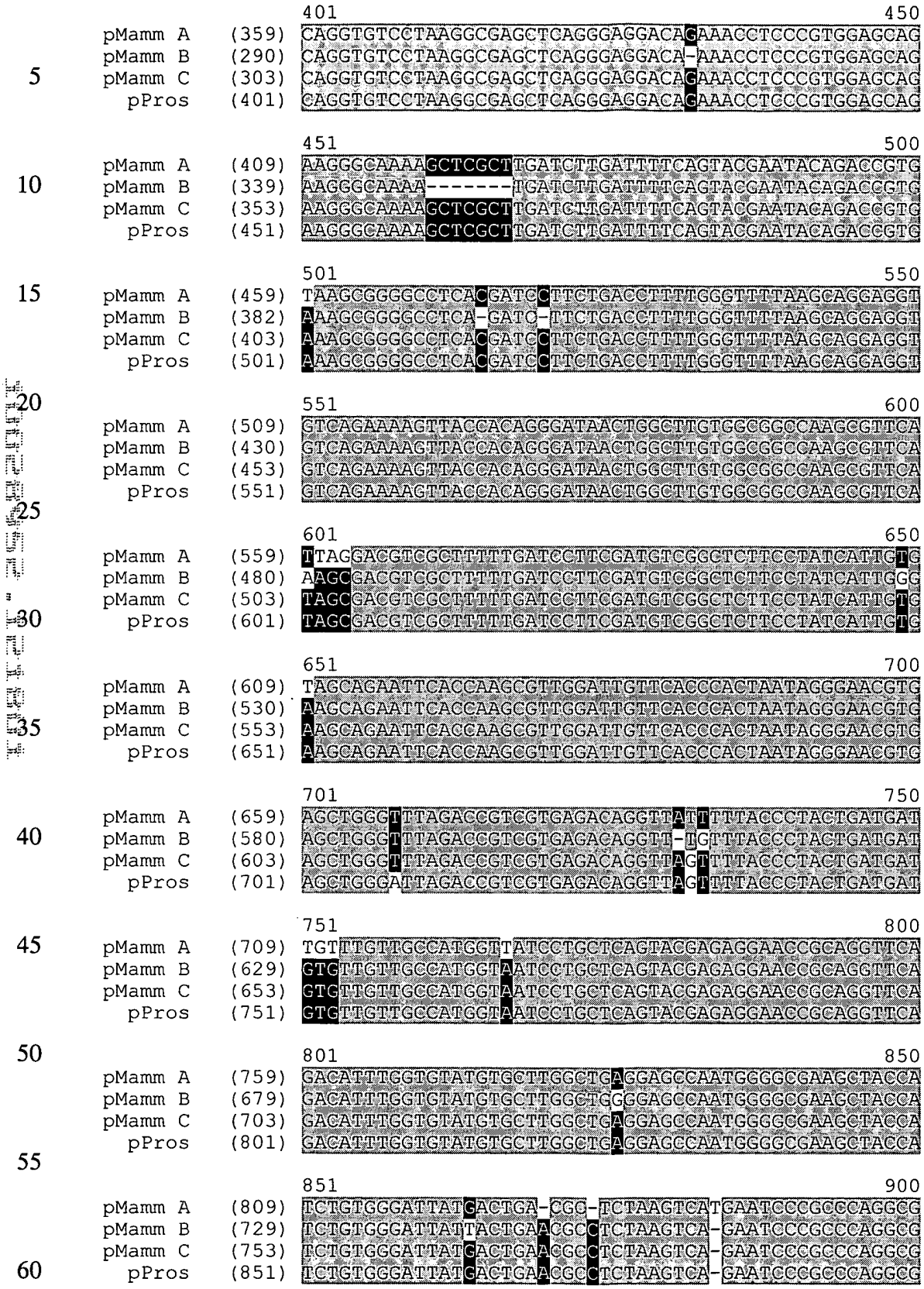
```
1851  ATCACACTCC AGATCTCTCC ATACCACACT GCTAGCACAT GTGCCTGTCA
      TAGTGTGAGG TCTAGAGAGG TATGGTGTGA CGATCGTGTA CACGGACAGT
      · L  I  P  G  S  C  Y  F  P  F  Y  F  L  S  L
1901  TCTTATTCCT GGCTCCTGTT ATTTCCCTTT TTATTTCCTT TCCCTTCCTC
      AGAATAAGGA CCGAGGACAA TAAAGGGAAA AATAAAGGAA AGGGAAGGAG
      · T  T  P  F  S  P  H  F  F  S  F  F  L  I  V
1951  CCACAACCCC TTTTTCCCCC CATTTCTTTT CTTTCTTTTT AATTGTTAAT
      GGTGTTGGGG AAAAAGGGGG GTAAAGAAAA GAAAGAAAAA TTAACAATTA
       Y  I  T  N  T  C  L  S  E  Q  L  I  *  H  K  R
2001  TACATAACTA ATACATGCTT ATCAGAACAA TTGATATAGC ACAAAAGGAT
      ATGTATTGAT TATGTACGAA TAGTCTTGTT AACTATATCG TGTTTTCCTA
      · *  S  T  G  E  *  *  L  I  P  V  I  L  A  L
2051  ATAAAGTACG GGTGAGTGAT AGCTCATCCC TGTAATCCTA GCACTTTGGA
      TATTTCATGC CCACTCACTA TCGAGTAGGG ACATTAGGAT CGTGAAACCT
      · A  K  A  G  R  S  L  E  S  R  V  R  D  Q  P
2101  AGGCCAAGGC AGGCAGATCA CTTGAGTCCA GAGTTCGAGA CCAGCCTGGG
      TCCGGTTCCG TCCGTCTAGT GAACTCAGGT CTCAAGCTCT GGTCGGACCC
       Q  H  G  E  T  L  S  L  Q  K  N  T  K  I  *  P
2151  CAACATGGTG AAACCCTGTC TCTACAAAAA AATACAAAAA TTTAGCCGGG
      GTTGTACCAC TTTGGGACAG AGATGTTTTT TTATGTTTTT AAATCGGCCC
      · V  L  A  H  T  C  S  L  S  Y  S  E  G  *  G
2201  CGTGCTGGCA CACACCTGTA GTCTCAGCTA CTCTGAGGGC TGAGGTGGGA
      GCACGACCGT GTGTGGACAT CAGAGTCGAT GAGACTCCCG ACTCCACCCT
      · I  D  *  A  Q  E  V  E  A  A  A  V  R  *  D
2251  AGATTGATTG AGCCCAGGAG GTGGAAGCTG CAGCAGTGCG CTGAGATTGC
      TCTAACTAAC TCGGGTCCTC CACCTTCGAC GTCGTCACGC GACTCTAACG
       A  I  A  L  Q  P  G  *  E  R  E  T  L  S  Q  K
2301  GCCATTGCAC TCCAGCCTGG GTGAGAGAGA GAGACCCTGT CTCAAAAAAA
      CGGTAACGTG AGGTCGGACC CACTCTCTCT CTCTGGGACA GAGTTTTTTT
      · K
2351  AAAAA
      TTTTT
```

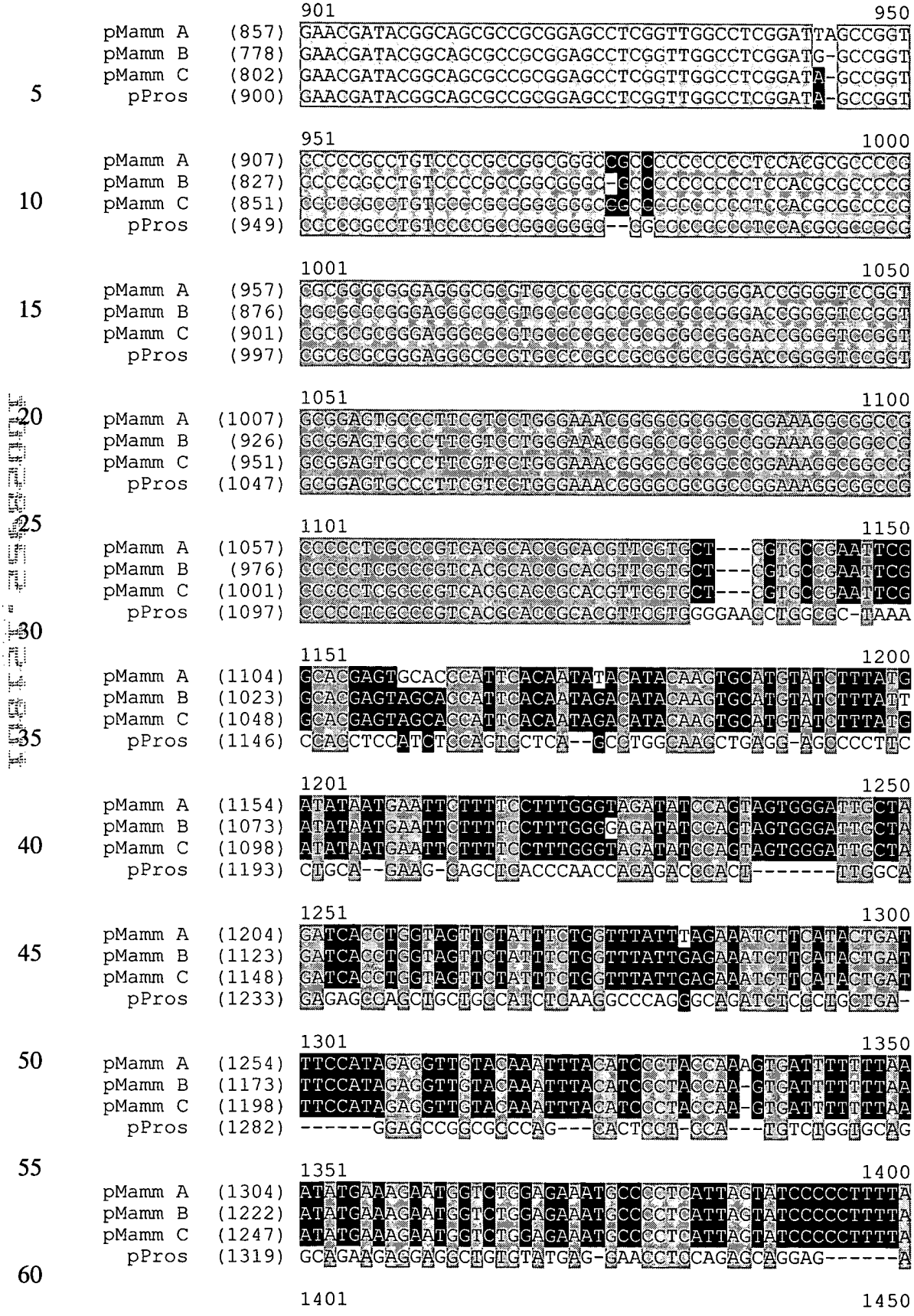
Table 5

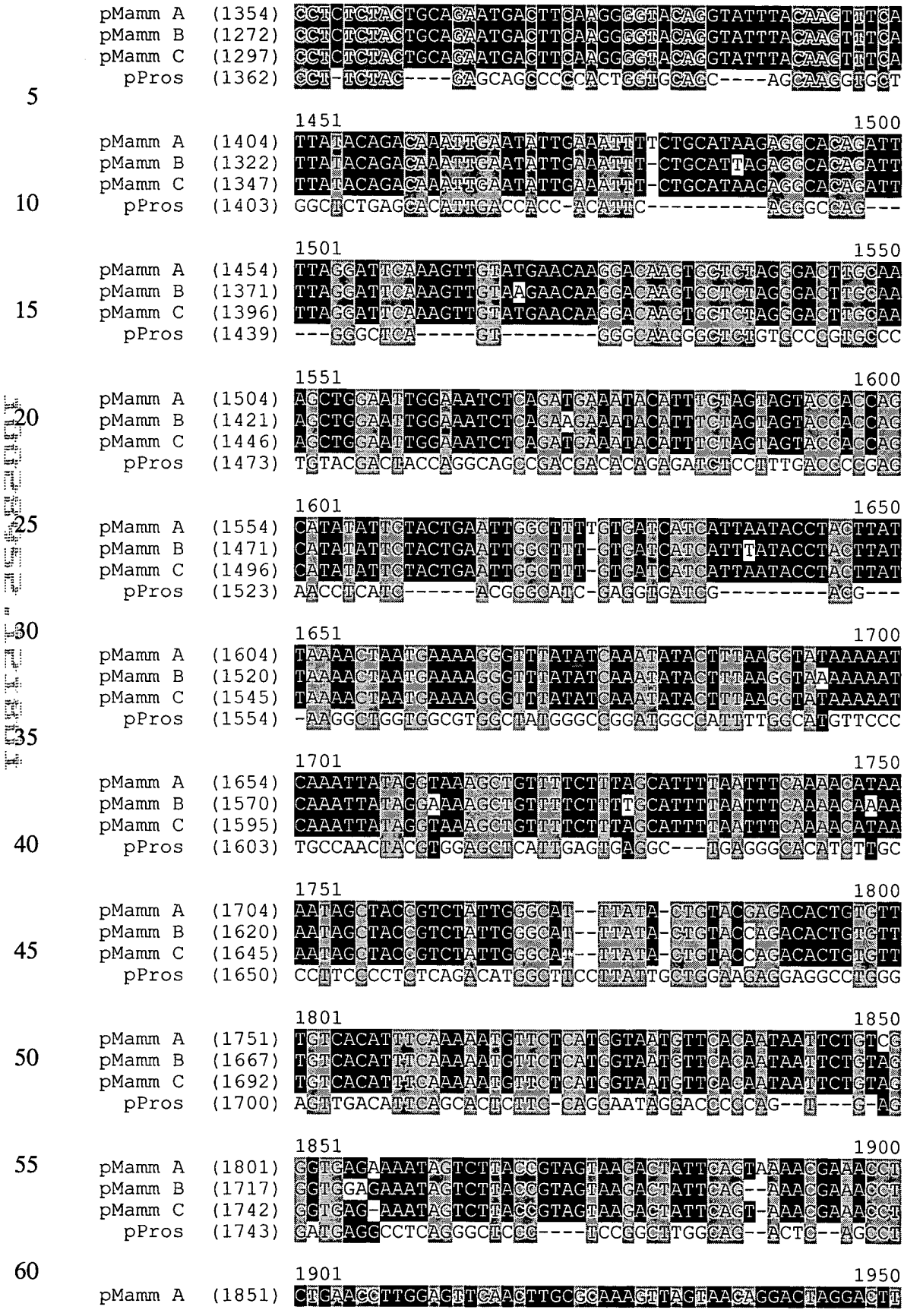
Comparison

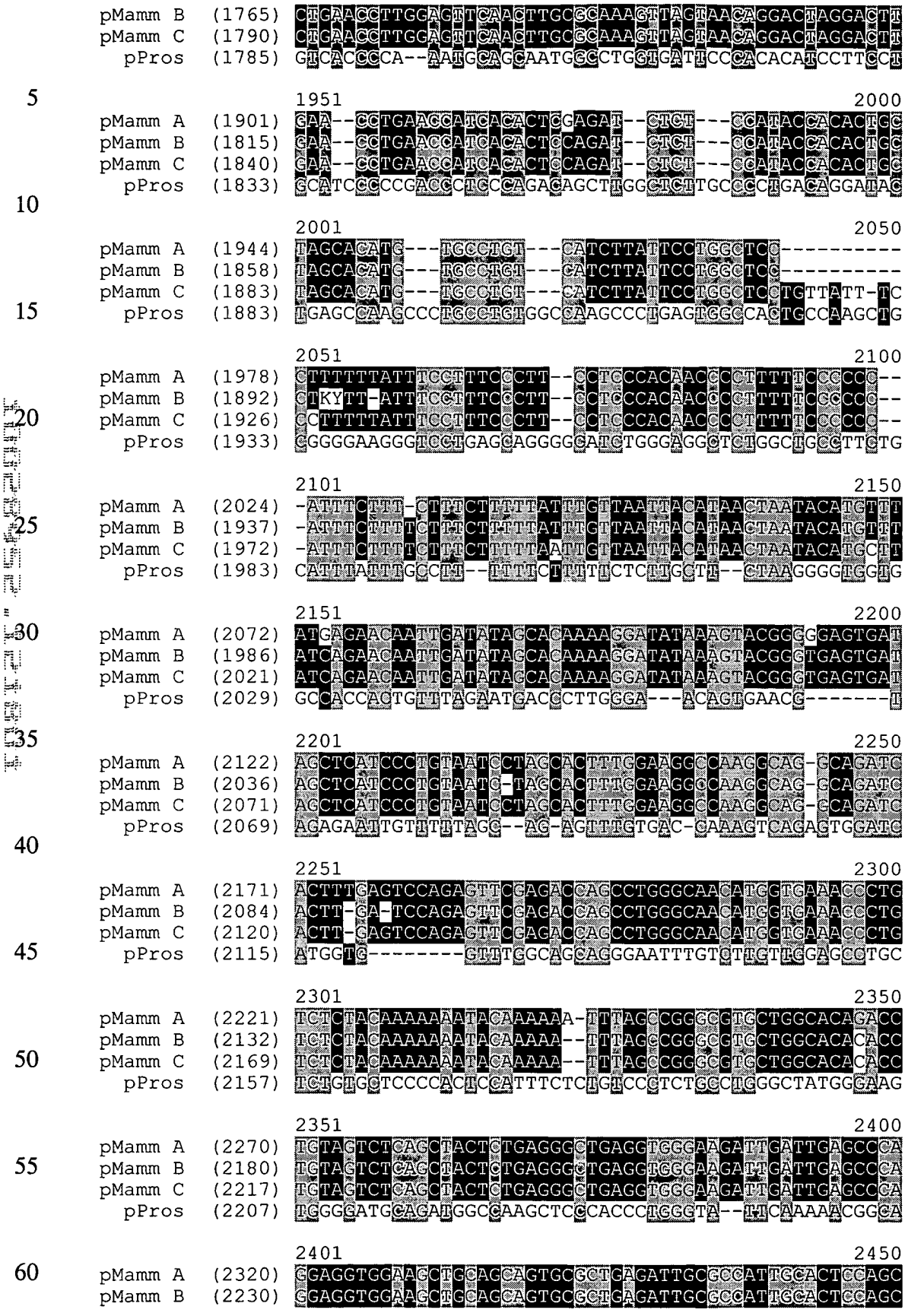
MammA, MammB, MammC, Prostate




401 450
pMamm A (359) CAGGTGTCCTAAGGCGAGCTCAGGGAGGACAGAAACCTCCCGTGGAGCAG
pMamm B (290) CAGGTGTCCTAAGGCGAGCTCAGGGAGGACA-AAACCTCCCGTGGAGCAG
pMamm C (303) CAGGTGTCCTAAGGCGAGCTCAGGGAGGACAGAAACCTCCCGTGGAGCAG
pPros (401) CAGGTGTCCTAAGGCGAGCTCAGGGAGGACAGAAACCTCCCGTGGAGCAG
451 500
pMamm A (409) AAGGGCAAAAGCTCGCTTGATCTTGATTTTCAGTACGAATACAGACCGTG
pMamm B (339) AAGGGCAAAA-------TGATCTTGATTTTCAGTACGAATACAGACCGTG
pMamm C (353) AAGGGCAAAAGCTCGCTTGATCTTGATTTTCAGTACGAATACAGACCGTG
pPros (451) AAGGGCAAAAGCTCGCTTGATCTTGATTTTCAGTACGAATACAGACCGTG
501 550
pMamm A (459) TAAGCGGGGCCTCACGATCCTTCTGACCTTTTGGGTTTTAAGCAGGAGGT
pMamm B (382) AAAGCGGGGCCTCA-GATC-TTCTGACCTTTTGGGTTTTAAGCAGGAGGT
pMamm C (403) AAAGCGGGGCCTCACGATCCTTCTGACCTTTTGGGTTTTAAGCAGGAGGT
pPros (501) AAAGCGGGGCCTCACGATCCTTCTGACCTTTTGGGTTTTAAGCAGGAGGT
551 600
pMamm A (509) GTCAGAAAAGTTACCACAGGGATAACTGGCTTGTGGCGGCCAAGCGTTCA
pMamm B (430) GTCAGAAAAGTTACCACAGGGATAACTGGCTTGTGGCGGCCAAGCGTTCA
pMamm C (453) GTCAGAAAAGTTACCACAGGGATAACTGGCTTGTGGCGGCCAAGCGTTCA
pPros (551) GTCAGAAAAGTTACCACAGGGATAACTGGCTTGTGGCGGCCAAGCGTTCA
601 650
pMamm A (559) TTAGGACGTCGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCATTGTG
pMamm B (480) AAGCGACGTCGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCATTGGG
pMamm C (503) TAGCGACGTCGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCATTGTG
pPros (601) TAGCGACGTCGCTTTTTGATCCTTCGATGTCGGCTCTTCCTATCATTGTG
651 700
pMamm A (609) TAGCAGAATTCACCAAGCGTTGGATTGTTCACCCACTAATAGGGAACGTG
pMamm B (530) AAGCAGAATTCACCAAGCGTTGGATTGTTCACCCACTAATAGGGAACGTG
pMamm C (553) AAGCAGAATTCACCAAGCGTTGGATTGTTCACCCACTAATAGGGAACGTG
pPros (651) AAGCAGAATTCACCAAGCGTTGGATTGTTCACCCACTAATAGGGAACGTG
701 750
pMamm A (659) AGCTGGGTTTAGACCGTCGTGAGACAGGTTATTTTACCCTACTGATGAT
pMamm B (580) AGCTGGGTTTAGACCGTCGTGAGACAGGTT-TGTTTACCCTACTGATGAT
pMamm C (603) AGCTGGGTTTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACTGATGAT
pPros (701) AGCTGGGATTAGACCGTCGTGAGACAGGTTAGTTTTACCCTACTGATGAT
751 800
pMamm A (709) TGTTTGTTGCCATGGTTATCCTGCTCAGTACGAGAGGAACCGCAGGTTCA
pMamm B (629) GTGTTGTTGCCATGGTAATCCTGCTCAGTACGAGAGGAACCGCAGGTTCA
pMamm C (653) GTGTTGTTGCCATGGTAATCCTGCTCAGTACGAGAGGAACCGCAGGTTCA
pPros (751) GTGTTGTTGCCATGGTAATCCTGCTCAGTACGAGAGGAACCGCAGGTTCA
801 850
pMamm A (759) GACATTTGGTGTATGTGCTTGGCTGAGGAGCCAATGGGGCGAAGCTACCA
pMamm B (679) GACATTTGGTGTATGTGCTTGGCTGGGGAGCCAATGGGGCGAAGCTACCA
pMamm C (703) GACATTTGGTGTATGTGCTTGGCTGAGGAGCCAATGGGGCGAAGCTACCA
pPros (801) GACATTTGGTGTATGTGCTTGGCTGAGGAGCCAATGGGGCGAAGCTACCA
851 900
pMamm A (809) TCTGTGGGATTATGACTGA-CGC-TCTAAGTCATGAATCCCGCCCAGGCG
pMamm B (729) TCTGTGGGATTATTACTGAACGCCTCTAAGTCA-GAATCCCGCCCAGGCG
pMamm C (753) TCTGTGGGATTATGACTGAACGCCTCTAAGTCA-GAATCCCGCCCAGGCG
pPros (851) TCTGTGGGATTATGACTGAACGCCTCTAAGTCA-GAATCCCGCCCAGGCG


901 950
pMamm A (857) GAACGATACGGCAGCGCCGCGGAGCCTCGGTTGGCCTCGGATTAGCCGGT
pMamm B (778) GAACGATACGGCAGCGCCGCGGAGCCTCGGTTGGCCTCGGATG-GCCGGT
pMamm C (802) GAACGATACGGCAGCGCCGCGGAGCCTCGGTTGGCCTCGGATA-GCCGGT
pPros (900) GAACGATACGGCAGCGCCGCGGAGCCTCGGTTGGCCTCGGATA-GCCGGT
951 1000
pMamm A (907) CCCCCGCCTGTCCCCGCCGGCGGGCCGCCCCCCCCCTCCACGCGCCCCG
pMamm B (827) CCCCCGCCTGTCCCCGCCGGCGGGC-GCCCCCCCCCTCCACGCGCCCCG
pMamm C (851) CCCCCGCCTGTCCCCGCCGGCGGGCGGCCGCCCCCCCTCCACGCGCCCCG
pPros (949) CCCCCGCCTGTCCCCGCCGGCGGGGGC--GGCCCCGCCGTCCACGCGCCCCG
1001 1050
pMamm A (957) CGCGCGCGGGAGGGCGCGTGCCCCGCCGCGCGCCGGGACCGGGGTCCGGT
pMamm B (876) CGCGCGCGGGAGGGCGCGTGCCCCGCCGCGCGCCGGGACCGGGGTCCGGT
pMamm C (901) CGCGCGCGGGAGGGCGCGTGCCCCGCCGCGCGCCGGGACCGGGGTCCGGT
pPros (997) CGCGCGCGGGAGGGCGCGTGCCCCGCCGCGCGCCGGGACCGGGGTCCGGT
1051 1100
pMamm A (1007) GCGGAGTGCCCTTCGTCCTGGGAAACGGGGCGCGGCCGGAAAGGCGGCCG
pMamm B (926) GCGGAGTGCCCTTCGTCCTGGGAAACGGGGCGCGGCCGGAAAGGCGGCCG
pMamm C (951) GCGGAGTGCCCTTCGTCCTGGGAAACGGGGCGCGGCCGGAAAGGCGGCCG
pPros (1047) GCGGAGTGCCCTTCGTCCTGGGAAACGGGGCGCGGCCGGAAAGGCGGCCG
1101 1150
pMamm A (1057) CCCCCTCGCCCGTCACGCACCGCACGTTCGTGCT---CGTGCCGAATTCG
pMamm B (976) CCCCCTCGCCCGTCACGCACCGCACGTTCGTGCT---CGTGCCGAATTCG
pMamm C (1001) CCCCCTCGCCCGTCACGCACCGCACGTTCGTGCT---CGTGCCGAATTCG
pPros (1097) CCCCCTCGCCCGTCACGCACCGCACGTTCGTGGGGAACCTGGCGC-TAAA
1151 1200
pMamm A (1104) GCACGAGTGCACCCATTCACAATATACATACAAGTGCATGTATCTTTATG
pMamm B (1023) GCACGAGTAGCACCATTCACAATAGACATACAAGTGCATGTATCTTTATT
pMamm C (1048) GCACGAGTAGCACCATTCACAATAGACATACAAGTGCATGTATCTTTATG
pPros (1146) CCACCTCCATCTCCAGTCCTCA--GCCTGGCAAGCTGAGG-AGCCCCTTC
1201 1250
pMamm A (1154) ATATAATGAATTCTTTTCCTTTGGGTAGATATCCAGTAGTGGGATTGCTA
pMamm B (1073) ATATAATGAATTCTTTTCCTTTGGGGAGATATCCAGTAGTGGGATTGCTA
pMamm C (1098) ATATAATGAATTCTTTTCCTTTGGGTAGATATCCAGTAGTGGGATTGCTA
pPros (1193) CTGCA--GAAG-CAGCTCACCCAACCAGAGACCCACT-------TTGGCA
1251 1300
pMamm A (1204) GATCACCTGGTAGTTCTATTTCTGGTTTATTTAGAAATCTTCATACTGAT
pMamm B (1123) GATCACCTGGTAGTTCTATTTCTGGTTTATTGAGAAATCTTCATACTGAT
pMamm C (1148) GATCACCTGGTAGTTCTATTTCTGGTTTATTGAGAAATCTTCATACTGAT
pPros (1233) GAGAGCCAGCTGCTGCCATCTCAAGGCCCAGGGCAGATCTCCCTGCTGA-
1301 1350
pMamm A (1254) TTCCATAGAGGTTGTACAAATTTACATCCCTACCAAAGTGATTTTTTTAA
pMamm B (1173) TTCCATAGAGGTTGTACAAATTTACATCCCTACCAA-GTGATTTTTTTAA
pMamm C (1198) TTCCATAGAGGTTGTACAAATTTACATCCCTACCAA-GTGATTTTTTTAA
pPros (1282) ------GGAGCCGGCGCCCAG---CACTCCT-CCA---TGTCTGGTGCAG
1351 1400
pMamm A (1304) ATATGAAAGAATGGTCTGGAGAAATGCCCCTCATTAGTATCCCCCTTTTA
pMamm B (1222) ATATGAAAGAATGGTCTGGAGAAATGCCCCTCATTAGTATCCCCCTTTTA
pMamm C (1247) ATATGAAAGAATGGTCTGGAGAAATGCCCCTCATTAGTATCCCCCTTTTA
pPros (1319) GCAGAAGAGGAGGCTGTGTATGAG-GAACCTCCAGAGCAGGAG------A
1401 1450


pMamm A (1354) CCTCTCTACTGCAGAATGACTTCAAGGGGTACAGGTATTTACAAGTTTCA
pMamm B (1272) CCTCTCTACTGCAGAATGACTTCAAGGGGTACAGGTATTTACAAGTTTCA
pMamm C (1297) CCTCTCTACTGCAGAATGACTTCAAGGGGTACAGGTATTTACAAGTTTCA
pPros (1362) CCT-TCTAC----GAGCAGCCCCCACTGGTGCAGC----AGCAAGGTGCT
1451 1500
pMamm A (1404) TTATACAGACAAATTGAATATTGAAATTTTCTGCATAAGAGGCACAGATT
pMamm B (1322) TTATACAGACAAATTGAATATTGAAATTT-CTGCATTAGAGGCACAGATT
pMamm C (1347) TTATACAGACAAATTGAATATTGAAATTT-CTGCATAAGAGGCACAGATT
pPros (1403) GGCTCTGAGGACATTGACCACC-ACATTC----------AGGGCCAG---
1501 1550
pMamm A (1454) TTAGGATTCAAAGTTGTATGAACAAGGACAAGTGGCTCTAGGGACTTGCAA
pMamm B (1371) TTAGGATTCAAAGTTGTAAGAACAAGGACAAGTGGCTCTAGGGACTTGCAA
pMamm C (1396) TTAGGATTCAAAGTTGTATGAACAAGGACAAGTGGCTCTAGGGACTTGCAA
pPros (1439) ---GGGCTCA-----GT--------GGGCAAGGGCTCTGTGCCCGTGCCC
1551 1600
pMamm A (1504) AGCTGGAATTGGAAATCTCAGATGAAATACATTTCTAGTAGTACCACCAG
pMamm B (1421) AGCTGGAATTGGAAATCTCAGAAGAAATACATTTCTAGTAGTACCACCAG
pMamm C (1446) AGCTGGAATTGGAAATCTCAGATGAAATACATTTCTAGTAGTACCACCAG
pPros (1473) TGTACGACTACCAGGCAGCCGACGACACAGAGATGTCCTTTGACCCCGAG
1601 1650
pMamm A (1554) CATATATTCTACTGAATTGGCTTTTGTGATCATCATTAATACCTAGTTAT
pMamm B (1471) CATATATTCTACTGAATTGGCTTT-GTGATCATCATTTATACCTAGTTAT
pMamm C (1496) CATATATTCTACTGAATTGGCTTT-GTGATCATCATTAATACCTACTTAT
pPros (1523) AACCTCATC------ACGGGCATC-GAGGTGATGG---------AGG---
1651 1700
pMamm A (1604) TAAAACTAATGAAAAGGGTTTATATCAAATATACTTTAAGGTATAAAAAT
pMamm B (1520) TAAAACTAATGAAAAGGGTTTATATCAAATATACTTTAAGGTAAAAAAT
pMamm C (1545) TAAAACTAATGAAAAGGGTTTATATCAAATATACTTTAAGGTATAAAAAT
pPros (1554) -AAGGCTGGTGGCGTGGCTATGGGCCGGATGGCCATTTTGGCATGTTCCC
1701 1750
pMamm A (1654) CAAATTATAGGTAAAGCTGTTTCTTTAGCATTTTAATTTCAAAAGATAA
pMamm B (1570) CAAATTATAGGAAAAGCTGTTTCTTTTGCATTTTAATTTCAAAACAAAA
pMamm C (1595) CAAATTATAGGTAAAGCTGTTTCTTTAGCATTTTAATTTCAAAACATAA
pPros (1603) TGCCAACTACGTGGAGCTCATTGAGTGAGGC---TGAGGGCACATGTTGC
1751 1800
pMamm A (1704) AATAGCTACCGTCTATTGGGCAT--TTATA-CTGTACGAGACACTGTGTT
pMamm B (1620) AATAGCTACCGTCTATTGGGCAT--TTATA-CTGTACCAGACACTGTGTT
pMamm C (1645) AATAGCTACCGTCTATTGGGCAT--TTATA-CTGTACCAGACACTGTGTT
pPros (1650) CCTTCCCCTCTCAGACATGGCTTCCTTATTGCTGGAAGAGGAGGCCTGGG
1801 1850
pMamm A (1751) TGTCACATTTCAAAAATGTTCTCATGGTAATGTTCACAATAATTCTGTCG
pMamm B (1667) TGTCACATTTCAAAAATGTTCTCATGGTAATGTTCACAATAATTCTGTAG
pMamm C (1692) TGTCACATTTCAAAAATGTTCTCATGGTAATGTTCACAATAATTCTGTAG
pPros (1700) AGTTGACATTCAGCACTCTTG-CAGGAATAGGACCCCCAG--T---G-AG
1851 1900
pMamm A (1801) GGTGAGAAAATAGTCTTACCGTAGTAAGACTATTCAGTAAAACGAAACCT
pMamm B (1717) GGTGGAGAAATAGTCTTACCGTAGTAAGACTATTCAG--AAACGAAACCT
pMamm C (1742) GGTGAG-AAATAGTCTTACCGTAGTAAGACTATTCAGT-AAACGAAACCT
pPros (1743) GATGAGGCCTCAGGGCTCCG----TCCGGCTTGGCAG--ACTG--AGCCT
1901 1950
pMamm A (1851) CTGAACGTTGGAGTTCAAGTTGCGGAAAGTTAGTAAGAGGACTAGGAGTT


pMamm B (1765) CTGAACCTTGGAGTTCAACTTGCGCAAAGTTAGTAACAGGACTAGGACTT
pMamm C (1790) CTGAACCTTGGAGTTCAACTTGCGCAAAGTTAGTAACAGGACTAGGACTT
pPros (1785) GTCACCCCA--AATGCAGCAATGGCCTGGTGATTCCCACACATCCTTCCT
1951 2000
pMamm A (1901) GAA--CCTGAAGCATCACACTCGAGAT--CTCT---CCATACCACACTGC
pMamm B (1815) GAA--CCTGAACCATCACACTCCAGAT--CTCT---CCATACCACACTGC
pMamm C (1840) GAA--CCTGAACCATCACACTCCAGAT--CTCT---CCATACCACACTGC
pPros (1833) GCATCCCCCGACCCTCCCAGACAGCTTGGCTCTTGCCCCTGACAGGATAC
2001 2050
pMamm A (1944) TAGCACATG---TGCCTGT---CATCTTATTCCTGGCTCC----------
pMamm B (1858) TAGCACATG---TGCCTGT---CATCTTATTCCTGGCTCC----------
pMamm C (1883) TAGCACATG---TGCCTGT---CATCTTATTCCTGGCTCCTGTTATT-TC
pPros (1883) TGAGCCAAGCCCTGCCTGTGGCCAAGCCCTGAGTGGCCACTGCCAAGCTG
2051 2100
pMamm A (1978) CTTTTTTATTTCCTTTCCCTT--CCTCCCACAACCCCTTTTTCCCCCC--
pMamm B (1892) CTKYTT-ATTTCCTTTCCCTT--CCTCCCACAACCCCTTTTTCCCCCC--
pMamm C (1926) CCTTTTTATTTCCTTTCCCTT--CCTCCCACAACCCCTTTTTCCCCCC--
pPros (1933) CGGGGAAGGGTCCTGAGCAGGGGCATCTGGGAGGCTCTGGCTGCCTTGTG
2101 2150
pMamm A (2024) -ATTTCTTT-CTTTCTTTTTATTTGTTAATTACATAACTAATACATGTTT
pMamm B (1937) -ATTTCTTTTCTTTCTTTTATTTGTTAATTACATAACTAATACATGTTT
pMamm C (1972) -ATTTCTTTTCTTTCTTTTAATTGTTAATTACATAACTAATACATGCTT
pPros (1983) CATTTATTTGCCTT--TTTTCTTTTTCTCTTGCTT--CTAAGGGGTGGTG
2151 2200
pMamm A (2072) ATGAGAACAATTGATATAGCACAAAAGGATATAAAGTACGGCGGAGTGAT
pMamm B (1986) ATCAGAACAATTGATATAGCACAAAAGGATATAAAGTACGGCTGAGTGAT
pMamm C (2021) ATCAGAACAATTGATATAGCACAAAAGGATATAAAGTACGGCTGAGTGAT
pPros (2029) GCCACCACTGTTTAGAATGACCCTTGGGA---ACAGTGAACG-------T
2201 2250
pMamm A (2122) AGCTCATCCCTGTAATCCTAGCACTTTGGAAGGCCAAGGCAG-GCAGATC
pMamm B (2036) AGCTCATCCCTGTAATG-TAGCACTTTGGAAGGCCAAGGCAG-GCAGATC
pMamm C (2071) AGCTCATCCCTGTAATCCTAGCACTTTGGAAGGCCAAGGCAG-GCAGATC
pPros (2069) AGAGAATTGTTTTTAGC--AG-AGTTTGTGAC-GAAAGTCAGAGTGGATC
2251 2300
pMamm A (2171) ACTTTGAGTCCAGAGTTCGAGACCAGCCTGGGCAACATGGTGAAACCCTG
pMamm B (2084) ACTT-GA-TCCAGAGTTCGAGACCAGCCTGGGCAACATGGTGAAACCCTG
pMamm C (2120) ACTT-GAGTCCAGAGTTCGAGACCAGCCTGGGCAACATGGTGAAACCCTG
pPros (2115) ATGGTG--------GTTTGGCAGCAGGGAATTTGTGTTGTTGGAGCCTGC
2301 2350
pMamm A (2221) TCTCTACAAAAAAATACAAAAA-TTTAGCCGGGCGTGCTGGCACAGACC
pMamm B (2132) TCTCTACAAAAAAATACAAAA--TTTAGCCGGGCGTGCTGGCACACACC
pMamm C (2169) TCTCTACAAAAAAATACAAAA--TTTAGCCGGGCGTGCTGGCACACACC
pPros (2157) TCTGTGGTCCCCACTCCATTTCTCTGTCCGTCTGCCTGGGCTATGGGAAG
2351 2400
pMamm A (2270) TGTAGTCTCAGCTACTCTGAGGGCTGAGGTGGGAAGATTGATTGAGCCCA
pMamm B (2180) TGTAGTCTCAGCTACTCTGAGGGCTGAGGTGGGAAGATTGATTGAGCCCA
pMamm C (2217) TGTAGTCTCAGCTACTCTGAGGGCTGAGGTGGGAAGATTGATTGAGCCCA
pPros (2207) TGGGGATGCAGATGGCCAAGCTCCCACCCTGGGTA--TTCAAAAACGGCA
2401 2450
pMamm A (2320) CGAGGTGGAAGCTGCAGCAGTGCGCTGAGATTGCGCCATTGCACTCCAGC
pMamm B (2230) CGAGGTGGAAGCTGCAGCAGTGCGCTGAGATTGCGCCATTGCACTCCAGC

```
pMamm C  (2267) GGAGGTGGAAGCTGCAGCAGTGCGCTGAGATTGCGCCATTGCACTCCAGC
  pPros  (2255) GACACAACATGTTCCTCCACGCGGCTCACTCGATGCC--TGCAGGCCCCA

                2451                                          2500
pMamm A  (2370) CTGGGTGAGAGAGAGAGACCCTGTCTCCAAAAAAAAAAAAAAAAAAAA-
pMamm B  (2280) CTGGGTGAGAGAGAGAGACCCTGTCTTCAAAAAAAAAAAAAAAAAAA---
pMamm C  (2317) CTGGGTGAGAGAGAGAGACCCTGTCTCAAAAAAAAAAAA----------
  pPros  (2303) GTGTGTGCCTCA-ACTGATTCTGACTTCAGGAAAAGTAAAAAAAAAAAA

                2501                          2532
pMamm A  (2419) --------------------------------
pMamm B  (2327) --------------------------------
pMamm C  (2356) --------------------------------
  pPros  (2352) AAAAACTCGAGAAGCTTTGGACTTCTTCGCCA
```